

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:13:22 ; Search time 70.93 Seconds

(without alignments)

111.842 Million cell updates/sec

Title: US-09-389-782A-1

Perfect score: 1263

Sequence: 1 EPKSCDKTHTCPPCPAPPELL.....MHEALHNHYTKSLSPGK 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36.*

1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1263	100.0	232	18 W26232	Human IgG1 hinge/F
2	1263	100.0	259	20 Y24154	Protein from pCds1
3	1263	100.0	329	17 R91806	Human immunoglobul
4	1263	100.0	351	14 R43685	Human kappa immuno
5	1263	100.0	371	10 P91918	Sequence of the li
6	1263	100.0	371	10 P93558	Linked human IgG
7	1263	100.0	376	19 W60037	Antigenic peptide
8	1263	100.0	379	19 W83962	Recombinant human
9	1263	100.0	379	19 W49073	Recombinant human
10	1263	100.0	396	18 W18574	Aggrecanase artifi
11	1263	100.0	396	18 W18575	Aggrecanase artifi
12	1263	100.0	400	21 Y15123	Porcine CTLA-4-Ig

13	1263	100.0	424	16 W14764	Human soluble kit
14	1263	100.0	424	16 W14765	Human soluble kit
15	1263	100.0	435	13 R26530	Sequence of one ch
16	1263	100.0	437	18 W10552	Alpha-1-acid glyco
17	1263	100.0	442	18 W10550	IgG1 polypeptide.
18	1263	100.0	445	20 Y24153	Bovine LOX-1 extra
19	1263	100.0	446	17 W05829	Humanised 1D10 ant
20	1263	100.0	447	20 Y31669	Human IgG1 chain C
21	1263	100.0	449	14 R43339	Completely humanis
22	1263	100.0	449	19 W49816	Amino acid sequenc
23	1263	100.0	452	20 Y30201	Heavy chain sequen
24	1263	100.0	459	14 R42066	Human anti-HBs hea
25	1263	100.0	460	21 Y69890	Human NR8a1pha/IgG
26	1263	100.0	461	14 R42162	Anti-HIV-1 recombi
27	1263	100.0	467	13 R22758	Reshaped C34 antib
28	1263	100.0	467	13 R22759	Reshaped C34 antib
29	1263	100.0	470	13 R22757	Reshaped C34 antib
30	1263	100.0	471	21 Y45030	HUMAN OCRI1-Fc fus
31	1263	100.0	475	13 R20057	Heavy chain of 3D6
32	1263	100.0	475	17 R93553	Monoclonal antibody
33	1263	100.0	475	18 W11641	Human anti-RSV mon
34	1263	100.0	475	18 W11639	Human anti-RSV mon
35	1263	100.0	476	14 R31023	Antibody D heavy c
36	1263	100.0	476	18 W01818	Primatized anti-hu
37	1263	100.0	476	18 W01822	Primatized anti-hu
38	1263	100.0	476	19 W63765	Macaque primatized
39	1263	100.0	476	19 W63761	Macaque primatized
40	1263	100.0	476	20 W88464	Monoclonal antibody
41	1263	100.0	477	20 W90207	hB7.2Fc soluble fu
42	1263	100.0	478	19 W63763	Macaque primatized
43	1263	100.0	480	20 W90206	hB7.1Fc soluble fu
44	1263	100.0	481	13 R24442	Sequence of antibo
45	1263	100.0	534	13 R26531	Sequence of CD4-Ig

ALIGNMENTS

RESULT 1

W26232 W26232 standard; Protein; 232 AA.

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX AC W26232;

XX A novel recombinant vector has been developed which comprises a
CC nucleotide sequence encoding a fusion protein. The fusion protein
CC comprises three domains joined together in order, from N-terminus to
CC C-terminus, of a first domain comprising a protein of interest, a second
CC domain comprising a hydrophilic spacer and an affinity domain, each
CC domain comprising amino acid residues. The present sequence represents
CC the hinge/Fc region of human IgG1, used in example 3 of the present
CC invention. The recombinant vector is used for the production of
CC authentic recombinant proteins of interest. The method of the invention
CC is useful for the expression of fusion proteins capable of isolation by
CC affinity chromatography in pro- or eukaryotic cells. This method allows
CC for the efficient cleavage and generation of authentic proteins of
CC interest that do not contain extraneous (i.e. non-naturally occurring)
CC amino acids.

XX Sequence 232 AA;

Query Match 100.0%; Score 1263; DB 18; Length 232;
Best Local Similarity 100.0%; Pred. No. 1e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EPKSCDKTHTCPPAPPELLGGPSVFLPFPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
Db 1 epksckdtkhtccpccpapelggpsvflfpkpkdtlmisrtpevtcvvvdvshedpevkf 60
Qy 61 NWYDGVGVHNAKTKPREEOYNSTYRVVSVLTVHLQDWLNKGKEYCKVSNKALPAPIKT 120
Db 61 nwydgvgevhnaaktprreeqynstyrvvsvltvhlqdwlngkeyckvsnkalpapiekt 120
Qy 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOGPENNYKTT 180
Db 121 iskakgqprepqvtytlppsrdeltnqvscltclvkgfypsdiavewesngogpennyyktt 180
Qy 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 232
Db 181 pvlsgdsfflyskltvdkstrwqgnvfscsvmhcalhnhytqkslsispkg 232

RESULT 2
Y24154
ID Y24154 standard; Protein; 259 AA.
AC Y24154;
DT 10-SEP-1999 (first entry)
DE Protein from pCd5lneg1 comprising human IgG1 Fc region genomic DNA.
KW LDL; denatured; oxidized; arteriosclerosis; hyperlipidaemia;
KW low density lipoprotein; receptor; detection; immunoglobulin;
KW fusion protein.
XX Homo sapiens.
OS Synthetic.
XX W09932520-A1.
PN 01-JUL-1999.
PD 18-DEC-1998; 98WO-05744.
XX 16-DEC-1998; 98JP-0358170.
PR 19-DEC-1997; 97JP-0364981.
PR 09-DEC-1998; 98JP-0349648.
XX (NISH) JAPAN TOBACCO INC.
PA Kakutani M, Masaki T, Sawamura T;
XX WPI; 1999-418906/35.
DR N-PSDB; X88533.

XX Fusion peptide for assay of oxidized LDL and for therapeutic use
PT Example 1; Page 92-96; 105pp; Japanese.
PS The present invention describes a fusion peptide which consists of the
XX extracellular domain of a mammalian oxidized LDL (low density
CC lipoprotein) receptor, fused to a partial heavy chain of a mammalian
CC immunoglobulin containing all or part of the constant region. Oxidized
CC LDL is a denatured form of LDL occurring in patients having
CC arteriosclerosis or hyperlipidaemia, and the fusion peptide can be
CC used for the assay of oxidized LDL in biological samples from such
CC patients, for the diagnosis of the disorders. It can also be used
CC therapeutically for the prevention and treatment of arteriosclerosis and
CC hyperlipidaemia. The present sequence represents the protein from the
CC vector DNA of pCd5lneg1 comprising human IgG1 Fc region genomic DNA.
XX Sequence 259 AA;

Query Match 100.0%; Score 1263; DB 20; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.2e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EPKSCDKTHTCPPAPPELLGGPSVFLPFPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
Db 28 epksckdtkhtccpccpapelggpsvflfpkpkdtlmisrtpevtcvvvdvshedpevkf 87
Qy 61 NWYDGVGVHNAKTKPREEOYNSTYRVVSVLTVHLQDWLNKGKEYCKVSNKALPAPIKT 120
Db 88 nwydgvgevhnaaktprreeqynstyrvvsvltvhlqdwlngkeyckvsnkalpapiekt 147
Qy 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOGPENNYKTT 180
Db 148 iskakgqprepqvtytlppsrdeltnqvscltclvkgfypsdiavewesngogpennyyktt 207
Qy 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 232
Db 208 pvlsgdsfflyskltvdkstrwqgnvfscsvmhcalhnhytqkslsispkg 259

RESULT 3
R91806
ID R91806 standard; Protein; 329 AA.
AC R91806;
XX 20-SEP-1996 (first entry)
DT Human immunoglobulin gamma heavy chain constant region sequences.
DE alkaline phosphatase; label; antibody; IgG; H-chain; C-region; CH1;
KW CH2; CH3; hinge; fusion protein; chimera; immunoassay.
XX Homo sapiens.
OS JP08070875-A.
XX 19-MAR-1996.
PD 05-SEP-1994; 94JP-0211035.
XX 05-SEP-1994; 94JP-0211035.
PR (TOYJ) TOSOH CORP.
XX WPI; 1996-203155/21.
DR N-PSDB; T27385.
XX Recombinant alkaline phosphatase (AP)-antibody fusion protein -
PT comprises AP fused downstream of antibody heavy or light chain,
PT useful as immunoassay reagent
XX

PS Example 1; Page 13-15; 44pp; Japanese.

XX The gene coding for human alkaline phosphatase is fused downstream
 CC of a gene coding for either the variable and CH1 regions of an
 CC antibody heavy chain or an antibody light chain. Coexpression of the
 CC H- and L-chain sequences, one of which is fused to the AP gene,
 CC results in production of AP-labelled antibodies suitable for use in
 CC immunoassays. The present sequence is from a human IgG heavy chain
 CC constant region.

XX Sequence 329 AA;

Query Match 100.0%; Score 1263; DB 17; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.6e-93;
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPCPAPPELLGPPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60
 Db 98 epksckdthtccppcpapellgppsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 157

QY 61 NWYVDGVEVHNKTPREQYNSTYRVSVTLVTHQDNLNGKEYCKVSNKALPAPIEKT 120
 Db 158 nwyvdgvevhnaktprreqynstyrsvsvtlvthqdwngkeyckvsnkalpapiekt 217

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEGSGQPENNYKTTTP 180
 Db 218 iskakgqprepqvtytlppsrdeltknqvsltcclvkgyfypsdiavevesngqpennykttt 277

QY 181 PVLDSGSPFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 232
 Db 278 pvltdsgsflyskltvdksrwggnvfscsvmhealhnhytqkslspsgk 329

RESULT 4
 ID R43685 standard; Protein: 351 AA.
 AC R43685;
 DT 25-MAY-1994 (first entry)

XX Human kappa immunoglobulin light chain constant domain.
 DE Human; immunoglobulin; constant; region; humanised; P-selectin; light;
 KW blocking; antibody; heavy; chain; variable; murine; thrombotic disease;
 KW monoclonal; PBI.3; CDR; complementarity determining region; leukocyte;
 KW expression vector; coexpression; pHCMV-1748RHA-gammalci-dhfr; epitope;
 KW pHCMV-1748RHA-KR-neo; PBI.3/Humanised version A; vascular endothelium;
 KW pHCMV-1747CH-gammaCi-neo; pHCMV-1747-CL-KR-neo; PBI.3 chimera;
 KW acute lung injury; ischaemia reperfusion injury; inflammation.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH Domain 22..119
 FT /note= "CH1 domain"
 FT Region 120..134
 FT /note= "Hinge region"
 FT Domain 135..244
 FT /note= "CH2 domain"
 FT Domain 245..352
 FT /note= "CH3 domain"

XX W09321956-A.
 PN 11-NOV-1993.
 XX 04-MAY-1993; 93WO-US04274.
 XX 05-MAY-1992; 92US-0880196.
 PR (CYTE-) CYTEL CORP.

XX Chestnut RW, Paulson JC, Polley MJ;
 PI WPI: 1993-368423/46.
 DR N-PSDB; Q51547.

XX Anti-P-selectin antibody for ischaemia acute lung injury treatment -
 PT useful to treat inflammation and pathological conditions of
 PT intercellular adhesion by competitive inhibition assays

XX Example 10; Fig 9; 82pp; English.

XX The sequences given in R43685-86 represent human immunoglobulin
 CC constant regions which were used in the production of the humanised
 CC P-selectin blocking antibody, along with the heavy and light chain
 CC variable region coding sequences of the murine monoclonal antibody
 CC PBI.3, given in R43687-88. The CDRs from PBI.3 heavy and light
 CC chains were substituted for the CDRs of human heavy and light chains.
 CC The humanised variable regions were inserted into expression vectors.
 CC By coexpression of appropriate combinations of heavy and light
 CC chains, several humanised antibodies can be expressed. Coexpression
 CC of pHCMV-1748RHA-gammalci-dhfr and pHCMV-1748RHA-KR-neo gives rise
 CC to the PBI.3/Humanised version A. Coexpression of pHCMV-1747CH-
 CC gammaCi-neo and pHCMV-1747-CL-KR-neo gives rise to the PBI.3 chimera.
 CC These humanised antibodies selectively bind epitopes on P-selectin and
 CC block adhesion of leukocytes to the vascular endothelium. They may be
 CC used to treat inflammatory and thrombotic diseases and other
 CC pathological conditions involving P-selectin and antibodies to it, esp.
 CC acute lung injury and ischaemia reperfusion injury.

XX Sequence 351 AA;

Query Match 100.0%; Score 1263; DB 14; Length 351;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPCPAPPELLGPPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60
 Db 120 epksckdthtccppcpapellgppsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 179

QY 61 NWYVDGVEVHNKTPREQYNSTYRVSVTLVTHQDNLNGKEYCKVSNKALPAPIEKT 120
 Db 180 nwyvdgvevhnaktprreqynstyrsvsvtlvthqdwngkeyckvsnkalpapiekt 239

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEGSGQPENNYKTTTP 180
 Db 240 iskakgqprepqvtytlppsrdeltknqvsltcclvkgyfypsdiavevesngqpennyktt 299

QY 181 PVLDSGSPFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 232
 Db 300 pvltdsgsflyskltvdksrwggnvfscsvmhealhnhytqkslspsgk 351

RESULT 5
 P91918
 ID P91918 standard; protein: 371 AA.
 AC P91918;
 DT 14-MAY-1990 (first entry)

XX Sequence of the linked immunoglobulin gamma chain constant region.
 KW Immunoglobulin gamma chain; IgG1 heavy chain constant region.
 XX Key Location/Qualifiers
 FH Misc-difference 42..43
 FT /note= "Insert site"
 FT Misc-difference 144..145
 FT /note= "Insert site"
 XX EP314317-A.

```

XX PD 03-MAY-1989.
XX PF 03-OCT-1988; 88EP-0309194.
XX PR 28-SEP-1988; 88US-0250785, US-104329.
XX PA (GETH ) GENENTECH INC.
XX PI Capon DJ, Gregory TJ;
XX PS WPI; 1989-131855/18.
XX DR N-PSDB; N90779.
XX CC Compsns. contg. ahdesion variants
XX PT - useful in therapy and diagnostics, eg CD4 variants
XX PT which are therapeutically useful for treating human
XX PT immuno-deficiency virus
XX PS Fig 4a-4b; ; 36pp; English.
XX CC It may be fused to the first 180 N-terminal
XX CC residues of CD4 at the C-terminus. The fusion protein may be used for
XX CC antiviral of immunomodulatory therapy particularly in treatment of HIV
XX CC infection.
XX SQ Sequence 371 AA;

Query Match 100.0%; Score 1263; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDSHEDPEVKF 60
DB 140 epkscdkhtcpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 199

QY 61 NWYDGVGVHNAKTPREEQYNSTYRVVSVLTVHLQDMLNGKEYCKVSNKALPAPIEKT 120
DB 200 nwydgvghvhnaktprreeqynstyrvvsvltvhlqdwlngkeyckvsnkalpapiekt 259

QY 121 ISKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTP 180
DB 260 iskakgpreqvtytlppsrdeltknqvsltcclvkgyfypsdiavwesngqpennykttp 319

QY 181 PVLDSGSGFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 232
DB 320 pvlsgsgfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispgk 371

RESULT 6
P93558
ID P93558 standard; protein; 371 AA.
XX AC P93558;
XX DT 06-JUN-1990 (first entry)
XX DE Linkered human IgG1 (gamma 1) chain fragment.
XX KW Human IgG1; gamma 1; immunoglobulin; CD4; fusion protein.
XX OS Homo sapiens.
XX PN W08902922-A.
XX PD 06-APR-1989.
XX PF 03-OCT-1988; 88WO-US03414.
XX PR 02-OCT-1987; 87US-0104329.
XX PA (GETH ) GENENTECH.

```

```

XX PI Capon DJ, Gregory TJ;
XX DR WPI; 1989-114397/15.
XX DR N-PSDB; N90736.
XX PT New nucleic acid sequences encoding adheson, esp. CD4, variants -
XX PT partic. with trans-membrane domain inactivated or fused to other
XX PT peptide, useful esp. for treating HIV infections
XX PS Fig 4A-4B2; pp. 10/13-12/13; 78pp; English.
XX CC It is employed in the prepn. of CD4 fusions. CD4 fusion proteins can
XX CC have antiviral and immunomodulatory activity and are esp. useful for
XX CC treating HIV infections, regardless of genetic variation within the
XX CC virus. They and antibodies raised against them can also be used
XX CC diagnostically for assaying adhesons and their ligands.
XX SQ Sequence 371 AA;

Query Match 100.0%; Score 1263; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDSHEDPEVKF 60
DB 140 epkscdkhtcpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 199

QY 61 NWYDGVGVHNAKTPREEQYNSTYRVVSVLTVHLQDMLNGKEYCKVSNKALPAPIEKT 120
DB 200 nwydgvghvhnaktprreeqynstyrvvsvltvhlqdwlngkeyckvsnkalpapiekt 259

QY 121 ISKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTP 180
DB 260 iskakgpreqvtytlppsrdeltknqvsltcclvkgyfypsdiavwesngqpennykttp 319

QY 181 PVLDSGSGFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 232
DB 320 pvlsgsgfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispgk 371

RESULT 7
W60037
ID W60037 standard; Protein; 376 AA.
XX AC W60037;
XX DT 11-SEP-1998 (first entry)
XX DE Antigenic peptide hFas (nd29) containing Fc region.
XX KW Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
XX KW heart failure; kidney failure; graft-versus-host disease; antibody;
XX KW myocardial infarction; ischemic restenosis; endotoxic shock.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..16
XX FT Protein /note= "hFas antigen signal peptide"
XX FT /note= "hFas (nd29) protein"
XX PN W09818487-A1.
XX PD 07-MAY-1998.
XX PF 31-OCT-1997; 97WO-JP03978.
XX PR 26-SEP-1997; 97JP-0262521.
XX PR 31-OCT-1996; 96JP-0290459.
XX PR 27-DEC-1996; 96JP-03511718.

```


XX (MOCH) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
XX Nagata S, Suda T, Yatomi T;
XX WPI; 1998-271925/24.
DR N-PSDB; V34430.
XX
XX Use of Fas antagonist for treatment and prevention of
PT apoptosis-related diseases - such as heart or kidney failure,
PT graft-versus-host disease or liver disease
XX
XX Examples; Fig 5-9; 86pp; Japanese.
XX
XX This represents the antigenic peptide hFas (nd29) containing the Fc
CC region. The invention provides the use of Fas antagonist as an agent for
CC the treatment and prevention of apoptosis-related diseases. The Fas
CC antagonist can be a partial Fas antigen peptide containing the
CC extracellular part of the protein, but lacking the signal sequence, an
CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is
CC preferably a humanised antibody. The Fas antagonist is used in the
CC treatment and prevention of diseases such as myocardial infarction, heart
CC failure, ischemic heart disease, acute kidney failure, graft-versus-host
CC disease, ischemic restenosis of the heart, liver or kidney, and
CC endotoxic shock, and also as an organ preservative in transplantation.
CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand
CC system.
XX
SQ Sequence 376 AA;

Query Match 100.0%; Score 1263; DB 19; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPKSCDKHTCTCPPELGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60
DB 145 epkscdkthtccpapeilggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 204
QY 61 NNYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120
DB 205 nnyvdgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiekt 264
QY 121 ISKAKGQPEPQVYTLPPSRDELTKNOVSLTCLVKGFPDSDIAVESNGQENNNYKTTTP 180
DB 265 iskakgqppepyvltlppsrdeltnqvsltdclvkgfypsdiavesngqennnykttcp 324
QY 181 PVLSDSGSFLLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 232
DB 325 pvlstdsgsfllyskltvdksrwqgnvfscsvmhealhnhytqkslspsgk 376

RESULT 8
ID W83962 standard; Protein; 379 AA.
XX AC W83962;
XX
DT 15-FEB-1999 (first entry)
XX Recombinant human metFc-OB protein.
XX
XX Recombinant; metFc-OB protein; Fc region; immunoglobulin; Ig; OB;
KW obesity; human; adiposity; blood lipid; diabetes type II; insulin;
KW hypoglycaemia; antihypertensive; diuretic; appetite suppressant;
KW suspension.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /note= "can be optionally replaced with Ala"

FT Misc-difference 20 /note= "can be optionally replaced with Glu"
FT Misc-difference 103 /note= "can be optionally replaced with Ala"
FT Misc-difference 105 /note= "can be optionally replaced with Ala"
FT Misc-difference 107 /note= "can be optionally replaced with Ala"
XX
PN W09846257-Al.
XX
PD 22-OCT-1998.
XX
XX 16-APR-1998; 98WO-US07828.
XX
PR 14-APR-1998; 98US-0059467.
PR 17-APR-1997; 97US-0843971.
XX
PA (AMGE-) AMGEN INC.
XX
XX Brenns DN, French DL, Speed MA;
XX
DR WPI; 1998-594525/50.
DR N-PSDB; V69685.
XX
XX Concentrated suspension of fusion of obesity protein with Fc
PT immunoglobulin fragment - stable at physiological pH, used for e.g.
PT reduction of weight and blood lipid levels, and for treatment of
PT type II diabetes
XX
PS Claim 2; Fig 5A-C; 47pp; English.
XX
XX This represents a recombinant metFc-OB protein which consists of an Fc
CC region of human immunoglobulin (Ig) fused to a human OB (obesity)
CC protein. The invention provides a human OB protein suspension that
CC contains at least 0.5 mg/ml of the human OB protein derivatised by
CC attachment of the Fc region of an Ig to the N-terminus of OB, and has a
CC pH 6-8. The suspensions are used to reduce weight, adiposity and blood
CC lipid levels, to treat or prevent diabetes type II, and to increase lean
CC mass and insulin sensitivity. They may be used in conjunction with
CC insulin, hypoglycaemics, antihypertensives, diuretics, appetite
CC suppressants etc. These suspensions are stable and active at
CC physiological pH and are ready-for-use formulations that do not require
CC freezing or freeze drying. As they are very concentrated, only small
CC volumes are required and they provide a sustained-release effect, with
CC increased potency and reduced frequency of injection.
XX
SQ Sequence 379 AA;

Query Match 100.0%; Score 1263; DB 19; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPKSCDKHTCTCPPELGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60
DB 2 epkscdkthtccpapeilggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 61
QY 61 NNYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120
DB 62 nnyvdgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiekt 121
QY 121 ISKAKGQPEPQVYTLPPSRDELTKNOVSLTCLVKGFPDSDIAVESNGQENNNYKTTTP 180
DB 122 iskakgqppepyvltlppsrdeltnqvsltdclvkgfypsdiavesngqennnykttcp 181
QY 181 PVLSDSGSFLLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 232
DB 182 pvlstdsgsfllyskltvdksrwqgnvfscsvmhealhnhytqkslspsgk 233
RESULT 9
W49073

ID W49073 standard; Protein; 379 AA.
XX W49073;
XX
XX
XX 18-NOV-1998 (first entry)
XX
XX Recombinant human MetFc-OB protein.
DE Recombinant human MetFc-OB protein.
XX
XX Recombinant human MetFc-OB protein; chimeric; immunoglobulin; diabetes;
KW high blood lipid level; arterial sclerosis; stroke; Fc-OB fusion protein.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX
XX Key Location/Qualifiers
FH Protein 2...379
FT /note= "Recombinant human Fc-OB protein"
FT Region 234...379
FT /note= "Human OB protein"
XX
XX WO9828427-A1.
XX
XX 02-JUL-1998.
XX
XX 11-DEC-1997; 97WO-US23183.
XX
XX 20-DEC-1996; 96US-0770973.
XX
XX (AMGE-) AMGEN INC.
XX
XX Hecht RI, Mann MB;
XX WPI; 1998-377658/32.
DR N-PSDB; V32900.
XX
XX New fusion proteins of OB and Fc - used for treating e.g. excess
PT weight, diabetes, arterial sclerosis, arterial plaque, high blood
PT lipid level, gall stones or stroke
XX
XX Claim 2; Fig 3A-3C; 107pp; English.
XX
XX The present sequence represents a recombinant human MetFc-OB fusion
CC protein. The invention provides Fc-OB fusion proteins whereby the
CC Fc region of an immunoglobulin or its analogue is linked, either directly
CC or indirectly using a linker, to the N-terminus of an OB protein or its
CC analogue. The Fc-OB fusion proteins are claimed to demonstrate increased
CC stability and clearance rate and decreased degradation as compared to OB
CC protein or a fusion of Fc to the C-terminus of the OB protein. These
CC Fc-OB fusion proteins are also claimed to be useful for treating excess
CC weight in an individual or animal or for treating co-morbidities
CC associated with excess fat such as diabetes, high blood lipid level,
CC arterial sclerosis and stroke.
XX
XX Sequence 379 AA;
SQ

Query Match 100.0%; Score 1263; DB 19; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPKSCDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
XX
DB 2 epkscdkthtccppcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 61
XX
QY 61 NWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120
XX
DB 62 NWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 121
XX
QY 121 ISKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTPP 180
XX
DB 122 ISKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTP 181
XX
QY 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFSVMHEALHNHYTQKSLSLSPCK 232
XX

Db 182 pvlsgdsfflyskltvdkswqgnvfscvmhealhnhytqkslsispck 233
XX
RESULT 10
W18574
ID W18574 standard; Protein; 396 AA.
XX
XX W18574;
AC
XX
XX 17-SEP-1997 (first entry)
XX
XX Aggrecanase artificial recombinant substrate ragg-1.
DE
XX
XX Artificial recombinant substrate; ragg1; aggrecanase; aggrecan;
KW osteoarthritis; diagnosis.
XX
XX Chimaeric Homo sapiens;
OS Chimaeric synthetic.
XX
XX
XX Key Location/Qualifiers
FH Peptide 1..24
FT /label= Sig-peptide
FT /note= "CD5 signal sequence"
FT Peptide 25..32
FT /label= FLAG
FT Domain 33..160
FT /label= Aggrecan
FT /note= "human aggrecan interglobular domain"
FT Peptide 161..164
FT /label= Spacer
FT Region 165..179
FT /label= Hinge
FT /note= "human IgG1 hinge region"
FT Region 180..289
FT /label= CH2
FT /note= "human IgG1 CH2 region"
FT Region 290..396
FT /label= CH3
FT /note= "human IgG1 CH3 region"
XX
XX EP785274-A1.
PN
XX
XX 23-JUL-1997.
PD
XX
XX 27-DEC-1996; 96EP-0120949.
PF
XX
XX 18-JAN-1996; 96EP-0100682.
PR
XX
XX (FARH) HOECHST AG.
PA
XX
XX Bartnik E, Buettner F, Caterson B, Eidenmueller B;
PI Hughes C;
PI
XX
XX WPI; 1997-365948/34.
DR
XX
XX N-PSDB; T69892.
DR
XX
XX Recombinant substrate for aggrecanase in vitro testing - and
PT encoding DNA, useful for studying aggrecanase activity e.g. by
PT detection of cleavage products for monitoring onset or progression
PT of osteoarthritis
XX
XX Claim 3; Page 15-16; 28pp; English.
PS
XX
XX An artificial recombinant substrate, ragg-1 (W18574), for
CC aggrecanase comprises the CD5 signal sequence, a FLAG epitope for
CC M1 monoclonal antibody detection, the interglobular domain of human
CC aggrecan, and human IgG1 hinge, CH2 and CH3 regions. It is the
CC expression product of a DNA molecule (T69892) that can be
CC incorporated into a vector for use in ragg-1 prodn. in host cells.
CC ragg-1 can be used in cell culture systems to study the activity of
CC aggrecanase, to detect new enzymatic cleavage sites, for the
CC affinity purification of aggrecanase, to isolate aggrecanase

CC cDNA by functional cloning, to screen for aggrecanase inhibitors,
CC in methods for monitoring the onset or progression of
CC osteoarthritis, and in diagnostic aids. Another rAGG-1 (W18575)
CC has alanine at amino acid position 34.
XX Sequence. 396 AA;
SQ

Query Match 100.0%; Score 1263; DB 18; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.9e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
DB 165 epksckdthctcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 224

QY 61 NMVVDGVEVHNKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120
DB 225 nmyvdgvevhnaktpreeqynstyrsvsvltvlhqdwlngkeyckvsnkalpapiekt 284

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 180
DB 285 iskakgqprepqvytlppsrdeltnqvsltcvkgyfypsdiavewesngqpennyykttp 344

QY 181 PVLDSGSPFLYSLKLTVDKSRMQQGNVFSCVMHEALHNHYTQKSLSLSPGK 232
DB 345 pvlsgsgsflyslkltvdksrwqgnvfscvmhealhnhytqkslsispgk 396

RESULT 11
W18575
ID W18575 standard; Protein; 396 AA.
XX
AC W18575;
XX
DT 17-SEP-1997 (first entry)
XX
DE Aggrecanase artificial recombinant substrate rAGG-1.
XX
KW Artificial recombinant substrate; rAGG1; aggrecanase; aggrecan;
KW Osteoarthritis; diagnosis.
XX
OS Chimaeric Homo sapiens;
OS Chimaeric synthetic.
FH
FT Key Location/Qualifiers
FT Peptide 1..24
FT /label= sig_peptide
FT /note= "CD5 signal sequence"
FT Peptide 25..32
FT /label= FLAG
FT Domain 33..160
FT /label= Aggrecan
FT /note= "human aggrecan interglobular domain"
FT Peptide 161..164
FT /label= Spacer
FT Region 165..179
FT /label= Hinge
FT /note= "human IgG1 hinge region"
FT Region 180..289
FT /label= CH2
FT /note= "human IgG1 CH2 region"
FT Region 290..396
FT /label= CH3
FT /note= "human IgG1 CH3 region"

EP785274-A1.
XX
PN
XX
PD
XX
PF 27-DEC-1996; 96EP-0120949.
XX
PR 18-JAN-1996; 96EP-0100682.

(FARH) HOECHST AG.
PA
XX Bartnik E, Buettner F, Caterson B, Eidenmueller B;
PI Hughes C;
XX
DR WPI; 1997-365948/34.
DR N-PSDB; T69893.
XX
XX Recombinant substrate for aggrecanase in vitro testing - and
PT encoding DNA, useful for studying aggrecanase activity e.g. by
PT detection of cleavage products for monitoring onset or progression
PT of osteoarthritis
XX
PS Claim 3; Refer to page 15-16; 28pp; English.
XX
CC An artificial recombinant substrate, rAGG-1 (W18575), for
CC aggrecanase comprises the CD5 signal sequence, a FLAG epitope for
CC M1 monoclonal antibody detection, the interglobular domain of human
CC aggrecan, and human IgG1 hinge, CH2 and CH3 regions. It is the
CC expression product of a DNA molecule (T69893) that can be
CC incorporated into a vector for use in rAGG-1 prodn. in host cells.
CC rAGG-1 can be used in cell culture systems to study the activity of
CC aggrecanase, to detect new enzymatic cleavage sites, for the
CC affinity purification of aggrecanase, to isolate aggrecanase
CC cDNA by functional cloning, to screen for aggrecanase inhibitors,
CC in methods for monitoring the onset or progression of
CC osteoarthritis, and in diagnostic aids. Another rAGG-1 (W18574)
CC has glycine at amino acid position 34.
XX
SQ Sequence 396 AA;

Query Match 100.0%; Score 1263; DB 18; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.9e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
DB 165 epksckdthctcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 224

QY 61 NMVVDGVEVHNKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120
DB 225 nmyvdgvevhnaktpreeqynstyrsvsvltvlhqdwlngkeyckvsnkalpapiekt 284

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 180
DB 285 iskakgqprepqvytlppsrdeltnqvsltcvkgyfypsdiavewesngqpennyykttp 344

QY 181 PVLDSGSPFLYSLKLTVDKSRMQQGNVFSCVMHEALHNHYTQKSLSLSPGK 232
DB 345 pvlsgsgsflyslkltvdksrwqgnvfscvmhealhnhytqkslsispgk 396

RESULT 12
Y15123
ID Y15123 standard; Protein; 400 AA.
XX
AC Y15123;
XX
DT 07-FEB-2000 (first entry)
XX
DE Porcine CTLA-4-Ig construct.
XX
KW Porcine CTLA-4; soluble protein; xenograft; organ transplant; B7; CD28;
KW xenograft-specific immunosuppression; recipient T-cell; anergy;
KW co-stimulatory signal 2; homology; human CTLA-4; bovine CTLA-4.
OS Sus scrofa.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Region 162..168

FT FT /label= Flexible_linker
 FT FT /note= "Denotes the junction between pCTLA-4"
 FT Domain 169..362
 FT FT /label= IgG1_domain

XX WO957266-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-GB01350.

XX 30-APR-1998; 98GB-0009280.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Lechler IR, Dorling A;

XX WPI; 2000-038815/03.

XX Inhibiting T-cell mediated rejection of xenotransplanted organs

XX Claim 1; Fig 4; 43pp; English.

CC The present sequence is porcine CTLA-4-Ig construct for xenograft
 CC -specific immunosuppression. In a pig-to-human transplantation, the
 CC soluble protein could comprise the extracellular domain of porcine CTLA-4
 CC fused to a human C gamma 1 chain of IgG1. This construct was subcloned
 CC into the expression vector pHOOK-3FM and used to transfect DAP.3 or
 CC CHO-K1 cells. pCTLA-4-Ig preferentially binds to porcine B7 and blocks
 CC its interaction with CD28 on recipient T-cells. This is useful as a
 CC species-specific reagent to inhibit human T-cell proliferative responses
 CC to a variety of stimulators.

XX Sequence 400 AA;

Query Match 100.0%; Score 1263; DB 21; Length 400;
 Best Local Similarity 100.0%; Pred. No. 2e-93;
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60

Db 169 epksckdthtcppcpapellggpsvflfppkpkdtlmisrtpevtcvvdvshedpevkf 228

QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLTQDQLNGKEYCKVSNKALPAPIEKT 120

Db 229 nwyvdgvevhnaktkpreeqynstyrvvsvltvltqgwlngkeyckvsnkalpapiekt 288

QY 121 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTP 180

Db 289 iskagqprepqvlytlppsrdelcknqvsclclvkgyfypsdiawesngopennykttt 348

QY 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPGK 232

Db 349 pvlsgdsfflyskltvdksrwggnvfscsvmhealhnhytqkslsispkg 400

RESULT 13

W14764

ID W14764 standard; Protein; 424 AA.

XX AC W14764;

XX 11-JUN-1997 (first entry)

XX Human soluble kit ligand-IgG fusion protein.

DE Kit ligand; c-kit proto-oncogene; cytokine; growth factor;
 KW haematopoietic cell; cell proliferation; stem cell; anaemia;
 KW thrombocytopaenia; therapy; IgG1.

XX Homo sapiens.

DE Human soluble kit ligand-IgG fusion protein (corrected).

XX Key Location/Qualifiers
 FH Peptide 1..25
 FT /label= Sig_peptide
 FT /note= "KL signal peptide"
 FT Protein 26..424
 FT FT /label= Mat_protein
 FT FT /note= "human KL-IgG fusion"

XX WO9526199-A1.

XX 05-OCT-1995.

XX 28-MAR-1995; 95WO-US03866.

XX 28-MAR-1994; 94US-0220379.

XX (CYTO-) CYTOMED INC.

XX Lobell RB, Nocka KH;

XX WPI; 1995-351198/45.

XX N-PSDB; T63109.

XX Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit
 XX increased activity in promoting cell proliferation

XX Claim 10; Page 43-44; 88pp; English.

CC A fusion protein (W14764) between human soluble kit ligand (KL)
 CC (see also W14761) and a human IgG1 heavy chain can be transiently
 CC expressed in COS cells transfected with a human KL-Ig cDNA
 CC construct (T63109) in vector CDM8; a corrected KL-Ig construct
 CC (W14765) has also been prepd. KL-Ig can also be produced as a
 CC dimer stabilised by intermolecular disulphide bonds or a peptide
 CC linker. The stabilised KL-Ig dimers have a more favorable cell
 CC proliferation:mast cell activation ratio than native KL and can
 CC stimulate haematopoietic recovery or stem cell/progenitor cell
 CC mobilisation with less toxicity.

XX Sequence 424 AA;

Query Match 100.0%; Score 1263; DB 16; Length 424;
 Best Local Similarity 100.0%; Pred. No. 2.1e-93;
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60

Db 193 epksckdthtcppcpapellggpsvflfppkpkdtlmisrtpevtcvvdvshedpevkf 252

QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLTQDQLNGKEYCKVSNKALPAPIEKT 120

Db 253 nwyvdgvevhnaktkpreeqynstyrvvsvltvltqgwlngkeyckvsnkalpapiekt 312

QY 121 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTP 180

Db 313 iskagqprepqvlytlppsrdelcknqvsclclvkgyfypsdiawesngopennykttt 372

QY 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPGK 232

Db 373 pvlsgdsfflyskltvdksrwggnvfscsvmhealhnhytqkslsispkg 424

RESULT 14

W14765

ID W14765 standard; Protein; 424 AA.

XX AC W14765;

XX 11-JUN-1997 (first entry)

XX Kit ligand; c-kit proto-oncogene; cytokine; growth factor;
 KW haematopoietic cell; cell proliferation; stem cell; anaemia;
 KW thrombocytopoiesis; therapy; IgG1.
 XX Homo sapiens.

Key Location/Qualifiers
 Peptide 1..25
 /label= Sig_peptide
 /note= "KL signal peptide"
 Protein 26..424
 /label= Mat_protein
 /note= "human KL-IgG fusion"

W09526199-A1.
 05-OCT-1995.
 28-MAR-1995; 95WO-US03866.
 28-MAR-1994; 94US-0220379.

(CYTO-) CYTOMED INC.
 Lobell RB, Nocka KH;
 WPI: 1995-351198/45.
 N-PSDB; T63110.

Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit
 increased activity in promoting cell proliferation
 Claim 10; Page 46-48; 88pp; English.

A fusion protein (W14765) between human soluble kit ligand (KL)
 (see also W14761) and a human IgG1 heavy chain can be transiently
 expressed in COS cells transfected with a human KL-Ig cDNA
 construct (T63110) in vector CDM8. KL-Ig can also be produced as a
 dimer stabilised by intermolecular disulphide bonds or a peptide
 linker. The stabilised KL-Ig dimers have a more favorable cell
 proliferation: mast cell activation ratio than native KL and can
 stimulate haematopoietic recovery or stem cell/progenitor cell
 mobilisation with less toxicity.

Sequence 424 AA;

Query Match 100.0%; Score 1263; DB 16; Length 424;
 Best Local Similarity 100.0%; Pred. No. 2.1e-93;
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
 Db 193 epkscdkthtccpcpapellgpgsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 252

QY 61 NWYDGVVHNKAKTPREEQYNSTYRVSVLTFLHODWLNKKEYCKVSNKALPAPIEKT 120
 Db 253 nwydgvvhnaktprreeqynstyrsvsvltflhdwlngkeyckvsnkalpapiekt 312

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 180
 Db 313 iskakgqprepqvtytlppsrdeltnqvsltcclvkgyfypsdiavevesngqpennyktt 372

QY 181 PVLDSGSEFLYSKLTVDKSRQQGQNVFSCSVNHEALHNHYTQKSLSLSPGK 232
 Db 373 pvltdsgsefllskltvdksrqgqnvfscsvnmhealhnhytqkslsispgk 424

RESULT 15
 R26530
 ID R26530 standard; Protein; 435 AA.
 XX

R26530;
 XX 28-JAN-1993 (first entry)
 DE Sequence of one chain of a CD4-gamma 1 chimeric heavy chain
 DE homodimer.
 XX CD4-gamma 1 chimeric heavy chain homodimer; expression vector; HIV;
 KW therapy; diagnostic agent; inhibition.
 XX Synthetic.

Key Location/Qualifiers
 Region 1..204
 /label= CD4
 /note= "1..25 = preregion"
 Region 205..219
 /label= hinge
 Region 220..329
 /label= CH2
 Region 330..436
 /label= CH3

W09213559-A.
 20-AUG-1992.
 10-FEB-1992; 92WO-US01152.
 08-FEB-1991; 91US-0654205.
 (PROG-) PROGENICS PHARM INC.

Beaudry GA, Maddon PJ;
 WPI: 1992-299758/36.
 N-PSDB; Q27830.

CD4-gamma 1 chimeric heavy chain homo-dimer and its expression
 vector - for preventing and treating HIV infection useful as a
 diagnostic agent

Example; Fig 3; 88pp; English.

Human CD4 cDNA was excised from pSP6T4 and cloned into M13mp18. The
 2 kb PstI/PstI fragment from pAR lambda 1 contg. the human lambda 1
 heavy chain gene (contg. the hinge, CH2 and CH3 exons) was isolated
 and cloned into the BAP-treated M13mp18/CD4 vector. To obtain a
 CD4-lambda 1 chimeric heavy chain gene, oligonucleotide-mediated
 site-directed mutagenesis was performed to juxtapose the CD4 and
 lambda 1 heavy chain DNA sequences, ligating the CD4 sequence in
 frame to the hinge exon. The DNA was then cloned into pCDNA-1 to
 produce CD4-IgG1-pcDNA1 (ATCC 40951).

Sequence 435 AA;

Query Match 100.0%; Score 1263; DB 13; Length 435;
 Best Local Similarity 100.0%; Pred. No. 2.2e-93;
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
 Db 204 epkscdkthtccpcpapellgpgsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 263

QY 61 NWYDGVVHNKAKTPREEQYNSTYRVSVLTFLHODWLNKKEYCKVSNKALPAPIEKT 120
 Db 264 nwydgvvhnaktprreeqynstyrsvsvltflhdwlngkeyckvsnkalpapiekt 323

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 180
 Db 324 iskakgqprepqvtytlppsrdeltnqvsltcclvkgyfypsdiavevesngqpennyktt 383

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:20:02 ; Search time 135.68 seconds
(without alignments)
346.406 Million cell updates/sec

Title: US-09-389-782A-2
Perfect score: 2179
Sequence: 1 MNKWLCCALLVLLDIIEWT.....OKLFLEMIGNQVSKISCL 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL15.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2179	100.0	401	11	O08712
2	2079	95.4	401	11	O08727
3	1892	86.8	401	4	O00300
4	1789	82.1	372	4	Q9UHP4
5	424.5	19.5	300	4	O95407
6	390.5	17.9	302	13	Q9PUS0
7	314	14.4	459	11	Q62327
8	302.5	13.9	482	11	O88734
9	301.5	13.8	655	4	O75509
10	298	13.7	439	4	Q16042
11	268.5	12.3	625	11	O35305
12	266.5	12.2	616	4	Q9Y606
13	226	10.4	355	12	O85308
14	224	10.3	349	12	O57098
15	222.5	10.2	349	12	O57305
16	221	10.1	349	12	O57284
17	220	10.1	349	12	O57097
18	219	10.1	349	12	O57110
19	218	10.0	349	12	O57111

20	218	10.0	349	12	Q89118
21	217.5	10.0	349	12	O57099
22	216	9.9	349	12	O89098
23	215.5	9.9	348	12	O57112
24	215.5	9.9	348	12	O85407
25	215.5	9.9	349	12	O57291
26	215.5	9.9	349	12	O57100
27	215.5	9.9	349	12	O57101
28	215.5	9.9	349	12	O57102
29	214.5	9.8	360	12	O57118
30	214	9.8	350	12	O57116
31	212.5	9.8	347	12	O57119
32	212	9.7	348	12	O57277
33	212	9.7	348	12	O57103
34	212	9.7	348	12	O57108
35	211.5	9.7	326	12	O57120
36	210.5	9.7	347	12	O57115
37	209.5	9.6	350	12	O57123
38	209	9.6	349	12	O57109
39	207.5	9.5	351	12	O57117
40	207.5	9.5	351	12	O73559
41	205.5	9.4	326	12	O57122
42	204.5	9.4	351	12	O57121
43	204	9.4	283	4	Q92956
44	204	9.4	283	4	Q9UM65
45	204	9.4	320	12	O57091

ALIGNMENTS

RESULT 1

O08712 ID O08712 PRELIMINARY; PRT; 401 AA.

AC O08712; 070202;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF).

GN TNFRSF11B OR OPB.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALE/C; TISSUE=KIDNEY;

RX MEDLINE=97262071; PubMed=9108485;

RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,

RA Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,

RA Derose M., Elliott R., Colombero A., Tan H.-L., Trill G., Sullivan J.,

RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,

RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,

RA Suggs S., Boyle W.J.;

RT "Osteoprotegerin: a novel secreted protein involved in the regulation

of bone density."

RL Cell 89:309-319(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129/OLA, AND NIH SWISS;

RX MEDLINE=98382527; PubMed=9714833;

RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,

RA Higashio K.;

RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)

gene and its expression in embryogenesis.";

RL Gene 215:339-343(1998).

CC -I- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES

OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY

SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO

OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN

STROMAL CELLS AND OSTEOCLAST PROGENITORS.

CC -I- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,
CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND
CC PLACENTA. NOT DETECTED IN SPLEEN.
CC -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT
CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY
CC 15 TO DAY 17.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94331; AAB53708.1; -.
DR EMBL; U94331; AAB53708.1; -.
DR EMBL; AB013898; BAA28269.1; -.
DR EMBL; AB013903; BAA33388.1; -.
DR EMBL; AB013899; BAA33388.1; JOINED.
DR EMBL; AB013900; BAA33388.1; JOINED.
DR EMBL; AB013901; BAA33388.1; JOINED.
DR EMBL; AB013902; BAA33388.1; JOINED.
DR HSSP; P25942; ICDF.
DR MGD; MGI:109587; Opg.
DR INTERPRO; IPR000488; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00117; DEATH DOMAIN; 1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 201 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT VARIANT 138 138 SWISS.
FT VARIANT 161 161 I -> R (IN STRAINS 129/OLA AND NIH
FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH
FT VARIANT 165 165 S -> A (IN STRAINS 129/OLA AND NIH
FT VARIANT 288 288 SWISS).
FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH
FT VARIANT 296 296 SWISS).
FT SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;
Query Match 100.0%; Score 2179; DB 11; Length 401;
Best Local Similarity 100.0%; Pred. No. 6.7e-164;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MNKWLCCALLVLLDIETWTQETLPKYLHYDPETGHQLLCKDCAPGYLKQHCYVRRKT 60
Db 1 MNKWLCCALLVLLDIETWTQETLPKYLHYDPETGHQLLCKDCAPGYLKQHCYVRRKT 60
Qy 61 LCVPCPDHSHYDTSWHTSDCYVCSVPVKELQSVKQECNTRNHRVCECEGRLYEFLCK 120
Db 61 LCVPCPDHSHYDTSWHTSDCYVCSVPVKELQSVKQECNTRNHRVCECEGRLYEFLCK 120
Qy 121 HRSPPGSGVVOAGTPERTVCKKCPDGFSGETSSKAPCIKHTNCSTFGLLLTQKGNAT 180
Db 121 HRSPPGSGVVOAGTPERTVCKKCPDGFSGETSSKAPCIKHTNCSTFGLLLTQKGNAT 180

Qy 181 HDNVCSGNREATQKCGIDVTLCCEAFFRFAVPTKTIIPNWLVLVDSLPCTKVNAESVERI 240
Db 181 HDNVCSGNREATQKCGIDVTLCCEAFFRFAVPTKTIIPNWLVLVDSLPCTKVNAESVERI 240
Qy 241 KRRHSSQQTQQLLKLWKHQRNDQEMVKIIQDIDLCSSVQRHLGHSNLTTEQLLALME 300
Db 241 KRRHSSQQTQQLLKLWKHQRNDQEMVKIIQDIDLCSSVQRHLGHSNLTTEQLLALME 300
Qy 301 SLPGKKISPEIERTRKTKCKSEQLLKLWLWRKNGQDQTLKGLMYALKHLKTSHPFKT 360
Db 301 SLPGKKISPEIERTRKTKCKSEQLLKLWLWRKNGQDQTLKGLMYALKHLKTSHPFKT 360
Qy 361 VTHSLRKTMRFLHSFTMYRLYQKLFLEMIGNOVQSVKISCL 401
Db 361 VTHSLRKTMRFLHSFTMYRLYQKLFLEMIGNOVQSVKISCL 401
RESULT 2
O08727 PRELIMINARY; PRT: 401 AA.
AC O08727;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
GN (OCIF).
DE TNFRSF11B OR OPG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
RN NCBI_TaxID=10116;
[1]
PC SEQUENCE FROM N.A.
RP TISSUE=INTESTINE;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,
Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
Davy E., Bucay N., Renshaw-Gagg L., Hughes T.M., Hill D., Pattison W.,
Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
of bone density.";
RL Cell 89:309-319(1997).
CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94330; AAB53707.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 201 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.


```

FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
SQ SEQUENCE 401 AA; 46192 MW; PEC6A31FD4E573A CRC64;

Query Match 95.4%; Score 2079; DB 11; Length 401;
Best Local Similarity 94.3%; Pred. No. 5e-156;
Matches 379; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MNKWLCCALLVLDIIETWTTQETLPPKYLHYDPETGHOLLCDKCAPGTYLKHQCTVRRKT 60
DB 1 MNKWLCCALLVLDIIETWTTQETLPPKYLHYDPETGHOLLCDKCAPGTYLKHQCTVRRKT 60

QY 61 LCVPCPDHSTDSWHTSDCVYSPVKELQSVKQECNRTHNRVCEEGRYLEIEFCLK 120
DB 61 LCVPCPDYSTDSWHTSDCVYSPVKELQSVKQECNRTHNRVCEEGRYLEIEFCLK 120

QY 121 HRCPPGSGVVGQATPERNTVCKPCDPGFFSGETSSKAPCKHTNCSLGLLIQKGNAT 180
DB 121 HRCPPGLGVLOQATPERNTVCKPCDPGFFSGETSSKAPCKHTNCSLGLLIQKGNAT 180

QY 181 HDNVCSGNREATKCGIDVTLCFAAFREAVPTKIIPNLVSLVDSLPSTKVNVAESVERI 240
DB 181 HDNVCSGNREATKCGIDVTLCFAAFREAVPTKIIPNLVSLVDSLPSTKVNVAESVERI 240

QY 241 KRRHSSQEQTFOLLKWKHQRNDQEMVKYIQQIDLCSSVQRHGLNSLTTEQLLALME 300
DB 241 KRRHSSQEQTFOLLKWKHQRNDQEMVKYIQQIDLCSSVQRHGLNSLTTEQLRLTME 300

QY 301 SLPGKKISPEIERTRKTCSSQLLKLKLSLWRKNGDQDTLGLMTALXHLKTSHPKPT 360
DB 301 SLPGKKISPEIERTRKTCSPQLLKLKLSLWRKNGDQDTLGLMTALXHLKAYHPKPT 360

QY 361 VTHSLRKTIRFLHSFTMYRYLQKLFLEMIGNQVQSVKISCL 401
DB 361 VTHSLRKTIRFLHSFTMYRYLQKLFLEMIGNQVQSVKISCL 401

RESULT 3
ID 000300 PRELIMINARY; PRT; 401 AA.
AC - 000300; 060236;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE OSTEOPTROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11B).
GN TNFRSF11B OR OPG OR OCIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=9262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.O., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density."
RL Call 89:309-319(1997).
RN [2]
RN SEQUENCE FROM N.A.
RP SEQUENCE=LUNG FIBROBLAST;

```

```

RX MEDLINE=98151033; PubMed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
RA Tsuda E., Morinaga T., Higashio K.;
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RT osteoclastogenesis in vitro."
RL Endocrinology 139:1329-1337(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; PubMed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human
RT osteoprotegerin/osteoclastogenesis-inhibitory factor."
RL Eur. J. Biochem. 254:685-691(1998).
CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; AB0021146; BAA25910.1; -
DR EMBL; AB008822; BAA32076.1; -
DR EMBL; AB008821; BAA32076.1; JOINED.
DR EMBL; U94332; BAB53709.1; -
DR HSSP; P25942; ICDF.
DR MIM; 602643; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 401 OSTEOPTROTEGERIN.
FT DOMAIN 23 183 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 86.8%; Score 1892; DB 4; Length 401;
Best Local Similarity 85.6%; Pred. No. 2.7e-141;
Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLDIIETWTTQETLPPKYLHYDPETGHOLLCDKCAPGTYLKHQCTVRRK 59
RN [1]
DB 1 MNKWLCCALLVLDIIETWTTQETLPPKYLHYDPETGHOLLCDKCAPGTYLKHQCTVRRK 59

```

```
Qy 60 TLCVPCPDHSTDSWHTSDECVYCSVKELQSVKQECNRTHNRVCECEGRLYEIEFCL 119
Db 60 TVCAPDHYVTDWHTSDECLYSPVKELQYVQECNRTHNRVCECKEGRYLEIEFCL 119
Qy 120 KHRSPGSGVVGAGTPERNVCKPCDGGFSGGETSSKAPCKHTNCSTFGLLLIQKQNA 179
Db 120 KHRSPGSGVVGAGTPERNVCKPCDGGFSGGETSSKAPCKHTNCSTFGLLLIQKQNA 179
Qy 180 THDNCVSGNREATOKCGIDVTLCCEAFPRFAVPTKIIFNWLSVLVDSPGKVNAAESVER 239
Db 180 THDNCVSGNREATOKCGIDVTLCCEAFPRFAVPTKIIFNWLSVLVDSPGKVNAAESVER 239
Qy 240 IKRRHSSQEQTFOLLKWLKHQNRQDMVKYKIIQDIDLCSSVQRHLGHSNLTTEQLLALM 299
Db 240 IKRRHSSQEQTFOLLKWLKHQNRQDMVKYKIIQDIDLCSSVQRHLGHSNLTTEQLLALM 299
Qy 300 ESLPGCKKISPEIERTRKTCSSBOLLKLLSLWRKNGDQDTLGLMYALKHLKTSHPPK 359
Db 300 ESLPGCKKISPEIERTRKTCSSBOLLKLLSLWRKNGDQDTLGLMYALKHLKTSHPPK 359
Qy 360 TVTHSLRKTMRFLHSFTMYRLYQKLFLEMIGNOVQSVKISCL 401
Db 360 TVTQSLKKTIRFLHSFTMYRLYQKLFLEMIGNOVQSVKISCL 401

RESULT 4
Q9UHP4 PRELIMINARY; PRT; 372 AA.
AC Q9UHP4
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE OSTEOPROTEGERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).
DR EMBL; AF134187; AAF20168.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
FT NON_TER 1
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 82.1%; Score 1789; DB 4; Length 372;
Best Local Similarity 86.0%; Pred. No. 3,2e-133;
Matches 320; Conservative 25; Mismatches 27; Indels 0; Gaps 0;

Qy 22 ETLPPKYLHYDPEYCHQLLCKAPGTYLKHQCVRRKTLVPCPDHSTYSWHTSDECV 81
Db 1 ETLPPKYLHYDEETSHQLLCKAPGTYLKHQCTAKWTKVCAPCPDHSTYSWHTSDECV 60

Qy 82 YCSVPCKELQSVKQECNRTHNRVCECEGRLYEIEFCLKHSRCPGSGVVGAGTPERNV 141
Db 61 YCSVPCKELQYVQECNRTHNRVCECKEGRYLEIEFCLKHSRCPGSGVVGAGTPERNV 120

Qy 142 CKKPDGFFSGTSSKAPCIKHTNCSTFGLLLIQKGNATHDNCVSGNREATOKCGIDVT 201
Db 121 CKRCPDGGFSSKAPCRKHTNCSTFGLLLIQKGNATHDNCVSGNREATOKCGIDVT 180

Qy 202 CEEAFFRFAVPTKIIPNWLVLVDSPGKVNAAESVERIKRHSQEQTFOLLKWLKHQ 261
Db 181 CEEAFFRFAVPTKIIPNWLVLVDSPGKVNAAESVERIKRHSQEQTFOLLKWLKHQ 240
```

```
Qy 262 RDQEMVKKIIQDIDLCSSVQRHLGHSNLTTEQLLALMESLPGRKISPEIERTRKTKCS 321
Db 241 KDQDIVKKIIQDIDLCSSVQRHLGHSNLTTEQLLALMESLPGRKISPEIERTRKTKCP 300
Qy 322 SEQLLKLLSLWRKNGDQDTLGLMYALKHLKTSHPKTYVTHSLRKTMRFLHSFTMYRLY 381
Db 301 SDQILKLLSLWRKNGDQDTLGLMYALKHLKTSHPKTYVTHSLRKTMRFLHSFTMYRLY 360
Qy 382 OKLFLEMIGNOV 393
Db 361 OKLFLEMIGNOV 372

RESULT 5
O95407 PRELIMINARY; PRT; 300 AA.
AC O95407
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN DCR3 OR TR6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99087326; PubMed=9872321;
RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA Goddard A.D., Botstein D., Ashkenazi A.;
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
RT colon cancer.";
RL Nature 396:699-703(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT "A newly identified member of tumor necrosis factor receptor
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL J. Biol. Chem. 274:13733-13736(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RX MEDLINE=20122600; PubMed=10655513;
RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RT cluster.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR EMBL; AF104419; AAD03056.1; -.
DR EMBL; AF134240; AAD29688.1; -.
DR EMBL; AF217796; AAF35244.1; -.
DR EMBL; AF217793; AAF33685.1; -.
DR EMBL; AF217794; AAF33686.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 19.5%; Score 424.5; DB 4; Length 300;
Best Local Similarity 39.0%; Pred. No. 7.1e-26;
```

Matches 80; Conservative 32; Mismatches 88; Indels 5; Gaps 2;

QY 26 PKYLHVDPTGHGKCDKCAPGTYLKQHCIVTRKTLVCPDHSYDTSWHTSDECVYCS 85
 Db 34 PTPWDAETGELVCAQCPGPFYVORCRRSPPTTCGCPGPPHYTFQFWNLYLRCRYCNV 93
 QY 86 VKELQSVKQECNRTNRVCECEGEGYLETEFLCKHRSCPPGSGVYVQAGTPERTVCKKC 145
 Db 94 LCGEREERACHATHNRACRCHTGFFAHAGFCLEHASCPGAGVAPGTPSQNTQCPC 153
 QY 146 PDGFFSGETSSKAPCIKHNCTSTFGLLLIQKGNATHDNV---CSGNREATQKCGIDVTLC 202
 Db 154 PPGTFSASSSSQCPHNRCTALGLANVPSSSHDTLCTSCGTPPLSTRVPGABE--C 211
 QY 203 EEAFFRFVPTKIIPNLSVLVDSL 227
 Db 212 ERAVIDFAQDISIKRLQRLQAL 236

RESULT 6
 QPUSO PRELIMINARY; PRT; 302 AA.

AC QPUSO;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE DECOY RECEPTOR
 OS Salvelinus fontinalis (Brook trout).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
 OX NCBI_TaxID=8038;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bobe J., Goetz F.W.;
 RT "A tumor necrosis factor receptor homolog is up-regulated in the brook trout (Salvelinus fontinalis) ovary at the completion of ovulation.";
 RL Biol. Reprod. 0:0-0(1999).
 DR EMBL; AF156738; AAD56428.1; -.
 DR HSSP; P19438; 1EXT
 DR INTERPRO; IPR000561; -.
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 1.
 KW Receptor.
 SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 17.9%; Score 390.5; DB 13; Length 302;
 Best Local Similarity 42.7%; Pred. No. 3.4e-23;
 Matches 67; Conservative 34; Mismatches 53; Indels 3; Gaps 2;

QY 32 DPETGHOLLCDKCAPGTYLKQHCIVTRKTLVCPDHSYDTSWHTSDECVYCSVPCKELQ 91
 Db 27 DRYSGLSIVCDRCPPGTYLAPCSAMRKSDCAECPNGAYTEFWNHISKCLRCS-MCAENQ 85
 QY 92 SVKQECNRTNRVCECEGEGYL--EIEFLCKHRSCPPGSGVYVQAGTPERTVCKKCPDGF 149
 Db 86 VVKQECSPSNCECEKEGYYFNKKYKACIKHKECPGPGYANTGTGTPHQDTECVQCAGF 145
 QY 150 FSGETSSKAPCIKHNCTSTFGLLLIQKGNATHDNVCS 186
 Db 146 YSEVSSAKATCLAQSNCKVGLRVKLGQDWHNTLCA 182

RESULT 7
 Q62327 PRELIMINARY; PRT; 459 AA.

AC Q62327;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=95178848; PubMed=7873884;
 RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
 RT "Allelic variation of the type 2 tumor necrosis factor receptor gene.";
 RL Mamm. Genome 5:726-727(1994).
 DR EMBL; X76401; CAA53981.1; -.
 DR HSSP; P19438; 1NCF.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 3.
 FT NON_TER 1
 FT VARIANT 87 87 S -> T.
 FT VARIANT 93 93 T -> I.
 FT VARIANT 268 268 F -> I.
 FT VARIANT 345 345 S -> F.
 FT VARIANT 421 421 Y -> C.
 SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match 14.4%; Score 314; DB 11; Length 459;
 Best Local Similarity 34.6%; Pred. No. 5.8e-17;
 Matches 66; Conservative 20; Mismatches 79; Indels 26; Gaps 4;

QY 18 WTTQETLPPKYL--HYDPETGH-----QLLCDKCAPGTYLKQHCIVTRKTL 62
 Db 2 WATGHTVPAQVLTTPKPEGYEQISQYDRKAQMCACPPGQYVHKFNKTSDTVC 61
 QY 63 VPCPDHSYDTSWHTSDECVYCSVPCKELQSVKQECNRTNRVCECEGEGYLEIEF----- 117
 Db 62 ADCEASMYTQVWNOFRCTCLSCSSCSDTDQVETRACTQQNRVCAACAGRYCALTHSGSC 121
 QY 118 --CLKHRSCPPGSGVYVQAGTPERTVCKKCPDGFSGFSGETSSKAPCIKHNCTSTFGLLLIQ 175
 Db 122 RQCMRLSKCGPGFGVASSRAPNGNLCKACAPGTFSDTSTSDVCRPHRCS----ILAI 177
 QY 176 KGNATHDNVCS 186
 Db 178 PGNASTDAVCA 188

RESULT 8
 O88734 PRELIMINARY; PRT; 482 AA.

AC O88734;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE P80 TNF-ALPHA RECEPTOR.
 GN TNFR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hurle B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;
 RT "The Mouse Tumor Necrosis Factor Receptor Gene:Genomic Structure and Characterization of the two Transcripts.";
 RL Genomics 0:0-0(0)
 DR EMBL; Y14619; CAA74969.1; -.
 DR EMBL; Y14620; CAA74969.1; JOINED.
 DR EMBL; Y14621; CAA74969.1; JOINED.
 DR EMBL; Y14622; CAA74969.1; JOINED.

Query Match 13.8%; Score 301.5; DB 4; Length 655;

Query Match 13.88; Score 301.5; DB 4; Length 655;

QY	31	YDPETH-----OLLCDKAPGYLKHOCNTRKTKLCVCPDHSYSDTSWHTSD	78
Db	10	YAPGSGTCRLREYYDOTAQCMCKSGPGHAKVFCIKTSDTVCDSCEDSTYYQLNNWP	69
QY	79	ECVYCSVCKELQSVKQECNRTNINVOECEBGRYLEI-----EFCCLKHRS CPPGSGVQ	132
Db	70	ECLSGSRCSDDQVETQACTREQNRICTCPGMYCALSKQEGRLCAPLKRCPGFVAR	129
QY	133	AGTPERNTVGKKPDGPFSGSETSSKAPCIKHTNCSTFGLLLIOGNATHDNVCS	186
Db	130	PGTSTDWCKPCAPGFTSNSTSDICRPHQICNVAI-----PGNASMDAVCT	179

QY	31	YDPETH-----OLLCDKAPGNYLKHCHVRRKTCVCPDHSYSDTSWHTSD	78
Db	10	YAPGSGTCRLREYYDOTAQCMCKSGPGHAKVFCIKTSOTVDCSDSEDSTYYQLNNWP	69
QY	79	ECVYCSVCKELQSVKQECNRTNINRVCEBEGYLEI-----EFLCLKHRSCTPPGSGVQ	132
Db	70	ECLSCGSRCSDDQVETQACTREQRNRTICRPGWYCALSKQEGRLCAPLKRCPGCVAR	129
QY	133	AGTPERNTVGKKCPDGFSGSETSSKAPCIKHTNCSTFGLLLIOGNATHDNVCS	186
Db	130	PGTSTSDWCKPCAPGFTSNSTSDICRPHQICNVAI-----PGNASMDAVCT	179

```

RESULT 11
O35305 PRELIMINARY; PRT; 625 AA.
O35305;
AC O35305;
TDT 01-JAN-1998 (TREMBlrel. 05, Created)
DDT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE RECEPTOR ACTIVATOR OF NF-KAPPA-B PRECURSOR (TNF-RELATED ACTIVATION-
DE INDUCED CYTOKINE RECEPTOR) (RANK).
DE TNFRSF1A OR RANK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER EPITHELIUM;
RR MEDLINE=98032977; PubMed=93671155;
RR Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
RR Tomatsko M.E., Roux E.R., Teepe M.C., DuBose R.F., Cosman D.,
RR Galibert L.;
RA "A homologue of the TNF receptor and its ligand enhance T-cell growth
RA and dendritic-cell function.";
RT Nature 390:173-179(1997).
RCC -|- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS RANKL.
RCC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
RCC -|- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
RDR EMBL; AF019046; BAB86810.1; -.
DR HSP; P25942; ICDF.
DR MGD; MGI:1314891; Tnfrsfla.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 1.
DR PRODOM; PD000711; -.
DR Receptor; Glycoprotein; Transmembrane; Signal; Repeat.
RCC SIGNAL 1 30
RCC CHAIN 31 625
RCC DOMAIN 31 214
RCC TRANSMEM 215 235
RCC DOMAIN 236 625
RCC DOMAIN 34 196
RCC REPEAT 34 70
RCC REPEAT 71 114
RCC REPEAT 115 153
RCC REPEAT 154 196
RCC REPEAT 35 47
RCC DISULFID 48 61
RCC DISULFID 51 69
RCC DISULFID 72 87
RCC DISULFID 93 113
RCC DISULFID 115 125
RCC DISULFID 127 134
RCC DISULFID 128 152
RCC DISULFID 155 170
RCC DISULFID 176 195
RCC CARBOHYD 106 106
RCC CARBOHYD 175 175
RCC SEQUENCE 625 AA; 66621 MW; F8C1872E99511D8E CRC64;
RCC 12.3%; Score 268.5; DB 11; Length 625;
RCC Best Local Similarity 29.9%; Pred. No. 3.2e-13;
RCC Matches 88; Conservative 34; Mismatches 123; Indels 49; Gaps 14;
RCC 7 CALLVLLDIEWTOETLPP--KYLHVDPTGTQLLCKDKCAPGTYLKQHCIVTRKTLCPV 64
RCC 19 CVLLVPLQV-----TLQVTPCTQERHYE-HLGR--CCSRCEPGKYLSSCKTPTSDSVCLP 71

```

FT DISULFID 175 194 BY SIMILARITY.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 616 AA; 56033 MW; E3DE9A7A08196F81 CRC64;

Query Match 12.2%; Score 266.5; DB 4; Length 616;
 Best Local Similarity 24.4%; Pred. No. 4.5e-13;
 Matches 108; Conservative 54; Mismatches 166; Indels 115; Gaps 21;

QY 5 LCCALLVLDIETWTTQETLPP--KYLHYDPTGTHQLCDKCAPGTYLKQHTVRRKILC 62
 DB 16 LLLALLARLQV-----ALQAPPTSEKHYE-HLGR--CCNKCEPKYMSSKCTTTSDSVC 68
 QY 63 VCPDHSYDTSWHTSDCVYCPVC---KELQSVKQECNTHNRVCECEEGRY--LEIEF 117
 DB 69 LFCGDEYLDWNEEDKCL-LHKVCDTGKALVAVAG-NSTTPRCACVAGVHNSQDEC 126
 QY 118 CLKHRSPPGSGVVOAGTPERNTVCKCPDGFSGTSSKAPCIKHTNCSFTGLLLIQKG 177
 DB 127 CRNTECAPGLAQHPQLQNKDCKPCLAGYFSDAFSSDKCRPTWNTCTFLGKRVEHGG 186
 QY 178 NATHNVCSGNREATQK-----CGDIVTLCEAFRFAVPTKII----- 216
 DB 187 TEKSDAVCVSSSLPARKPNPNEPHVLPGLILL---LFASVALVAIIIFGVCYRKKGKALT 243
 QY 217 ---PNWLSVLVSLPCTKVNAESVERIKRRHSS-----OEOTFOLLKLW 257
 DB 244 ANLWHINEACRLSGDK--ESSGSCSVSTHTANFGQOGACBGLVLLLEKTFPDMCY 301
 QY 258 KH-----QNRDQEMVKKIIQDIDLCSSVOR-----HLGHSNLTTEQL 295
 DB 302 PDQGGVCGCTCVGGPYAGGEDARML-SLVSKTEIEEDSFQMPTEDEYMDRPSQPTDQL 360
 QY 296 LALMESLPGKKISP--BEIE-----RTRKT-----KXSSQOLLKLLSLWRK 335
 DB 361 LFLTE--FGSKSTPFSEPLEYGENDSLSCQFTGTQSTVGSESCNCTEPLCRT----- 411
 QY 336 NGDQDTLGLMVALKHLKTSHP 358
 DB 412 --DWTPTMSENVLQEVDSGHCP 432

RESULT 13
 Q85308 PRELIMINARY; PRT; 355 AA.
 AC Q85308;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE SECRETED RECEPTOR BINDING TUMOR NECROSIS FACTOR (CRMB).
 GN CRMB.
 OS Cowpox virus (CPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10243;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRIGHTON RED;
 RX MEDLINE=83117629; PubMed=6961398;
 RA Pickup D.J., Bastia D., Stone H.O., Joklik W.K.;
 RT "Sequence of terminal regions of cowpox virus DNA: arrangement of repeated and unique sequence elements."
 RL Proc. Natl. Acad. Sci. U.S.A. 79:7112-7116(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRIGHTON RED;
 RX MEDLINE=90177240; PubMed=2309453;
 RA Parsons B.L., Pickup D.J.;
 RT "Transcription of orthopoxvirus telomeres at late times during infection."
 RL Virology 175:69-80(1990).
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN-BRIGHTON RED;
 RX MEDLINE=91196263; PubMed=2014645;
 RA Hu F.Q., Pickup D.J.;
 RT "Transcription of the terminal loop region of vaccinia virus DNA is initiated from the telomere sequences directing DNA resolution."
 RL Virology 181:716-720(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRIGHTON RED;
 RX MEDLINE=94378510; PubMed=8091665;
 RA Hu F.Q., Smith C.A., Pickup D.J.;
 RT "Cowpox virus contains two copies of an early gene encoding a soluble secreted form of the type II TNF receptor."
 RL Virology 204:343-356(1994).
 DR EMBL; L08906; ARA60952.1; .
 DR HSSP; P19438; IEXT.
 DR INTERPRO; IPR001368; .
 DR PFAM; PF00020; TNFR_C6; 2.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
 DR PROSITE; PS00500; TNFR_NGFR_2; 1.
 DR PRODOM; PD000771; .
 SQ SEQUENCE 355 AA; 39008 MW; 2C9E5C0D42FA4B3 CRC64;

Query Match 10.4%; Score 226; DB 12; Length 355;
 Best Local Similarity 24.1%; Pred. No. 3.7e-10;
 Matches 85; Conservative 49; Mismatches 141; Indels 78; Gaps 15;

QY 9 LVLVLDIETWTTQETLPPKYLHYDPTG-----HQLCDKCAPGTYLKQHC--TVR 57
 DB 6 LLLLLLSCIIIIISDITP-----HEPSNGKCKDNEYKRHLCLCLSPGPGTYASRLCDSKTN 60
 QY 58 RKTLCVPCPDHSYDTSWHTSDCVYCPVCRELQSVKQECNTHNRVCECEGRYLEI-- 115
 DB 61 TMTQCTPCASDTFTSRNNHLPACLSGRCDSNQVETRSCNTHNRICDCAPGYTCFLKG 120
 QY 116 ----EFLKLRHSCPPGSGVVOAGTPERNTVCKCPDGFSGTSSKAPCIKHTNCSFTGL 171
 DB 121 SSGCRACVSOYKCGIGYG-VSGHTTPTGDVCSPCGLGTYSHTVSSVDKC-EPVPSNTFNY 178
 QY 172 LLIQKG-NATHDNVC-----SGNREATOKGIDVTL---CEEAFRFAVPTKIIPNMLS 221
 DB 179 IDVEINLPVNDTSCRTTTTGLSEISITSELTITMNHKDCDPVFRN-----GVFS 229
 QY 222 VLVDLSLG-----TKV---NAESVERIKRRSSQEQTFQLLKLWKHQRDOEM 266
 DB 230 VLNEVATSGFTGQNRQYONISKVCTLNFEIKCNKNDKSYSSSKQ-----LTKTKNDOSI 283
 QY 267 V--KKIIQDIDLCSSVQRHLGHSNLTTE-----QLLALMESLPGK 305
 DB 284 MPHSESVTLVGDCLSVDIYILYSNTQDYETDTSIVHGVNVLVDVDSHMPGR 336

RESULT 14
 Q57098 PRELIMINARY; PRT; 349 AA.
 ID Q57098;
 AC Q57098;
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS Camelpox virus (strain CP-1).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=28873;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SAUDI-M3;
 RA Loparev V.N., Parsons J.M., Esposito J.J.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U87839; AAB94356.1; .

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:16:58 ; Search time 70.93 Seconds
(without alignments)
193.313 Million cell updates/sec

Title: US-09-389-782A-2
Perfect score: 2179
Sequence: 1 MNKWLCCALLVLLDIEWT.....QKLFLEMIGNQVSKISCL 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2179	100.0	401	18 W38344	Mouse osteoprotegerin
2	2079	95.4	401	18 W38343	Rat osteoprotegerin
3	1900	87.2	401	20 Y05742	Tumour necrosis factor
4	1900	87.2	401	20 W95030	Tumour necrosis factor
5	1900	87.2	401	20 W83926	Human FTHMA-070 pr
6	1895	87.0	401	17 R99925	Full length osteoc
7	1895	87.0	401	17 W53239	Human OCIF genome
8	1892	86.8	401	18 W38345	Human osteoprotegerin
9	1892	86.8	401	21 Y43400	Osteoprotegerin pro
10	1891	86.8	401	19 W57635	TR1 receptor prote
11	1885	86.5	401	17 R99932	Mutated OCIF, OCIF
12	1885	86.5	401	17 R99931	Mutated OCIF, OCIF

13	1882	86.4	399	17 R99942	Mutated OCIF, OCIF
14	1880	86.3	401	17 R99934	Mutated OCIF, OCIF
15	1879	86.2	401	17 R99933	Mutated OCIF, OCIF
16	1875	86.0	401	17 R99935	Mutated OCIF, OCIF
17	1862.5	85.5	391	19 W53238	Human OCIF genome
18	1852	85.0	395	19 W57636	Modified TR1 recep
19	1849	84.9	393	17 R99948	Mutated OCIF, OCIF
20	1828	83.9	380	17 R99924	Mature osteoclasto
21	1702.5	78.1	390	17 R99357	Human tumour necro
22	1670.5	76.7	360	17 R99936	Mutated OCIF, OCIF
23	1667	76.5	351	17 R99943	Mutated OCIF, OCIF
24	1645	75.5	359	17 R99939	Mutated OCIF, OCIF
25	1645	75.5	360	17 R99938	Mutated OCIF, OCIF
26	1620	74.3	359	17 R99937	Mutated OCIF, OCIF
27	1539	70.6	327	17 R99941	Mutated OCIF, OCIF
28	1510	69.3	321	17 R99949	Mutated OCIF, OCIF
29	1486.5	68.2	326	17 R99940	Mutated OCIF, OCIF
30	1481.5	68.0	349	20 W83928	Human FTHMA-070 pr
31	1334	61.2	272	17 R99944	Mutated OCIF, OCIF
32	1096	50.3	420	20 W89224	Tumour necrosis fa
33	1075.5	49.4	417	20 W89225	Tumour necrosis fa
34	1073	49.2	208	20 W89231	Mouse osteoprotege
35	1029	47.2	208	20 W89230	Rat osteoprotegeri
36	973	44.7	197	17 R99945	Mutated OCIF, OCIF
37	971.5	44.6	397	20 W89227	Tumour necrosis fa
38	946	43.4	208	20 W89232	Human osteoprotege
39	933	42.8	187	21 Y77464	Primate protein
40	928	42.6	187	17 R99950	Mutated OCIF, OCIF
41	810	37.2	366	20 W89228	Tumour necrosis fa
42	709	32.5	143	17 R99946	Mutated OCIF, OCIF
43	666.5	30.6	145	17 R99930	Osteoclastogenesis
44	665	30.5	154	17 R99929	Osteoclastogenesis
45	536.5	24.6	311	20 W89229	Tumour necrosis fa

ALIGNMENTS

RESULT	1
W38344	W38344 standard; Protein; 401 AA.
ID	XX
AC	XX
W38344;	XX
DT	20-APR-1998 (first entry)
XX	XX
DE	Mouse osteoprotegerin.
XX	XX
KW	Osteoprotegerin; antibody; diagnosis; affinity purification;
KW	recombinant production; transgenic animal; treatment; prevention;
KW	antisense oligonucleotide; probe; detection; screening; mouse;
KW	bone disease; osteoporosis; Paget's disease; hypercalcaemia;
KW	hyperparathyroidism; rheumatoid arthritis; osteomyelitis;
KW	osteolytic metastasis; periodontal bone loss; bone necrosis;
XX	XX
OS	Mus sp.
XX	XX
Key	Location/Qualifiers
CDS	90..1295
FT	/*tag= a
FT	/product= osteoprotegerin
DE19654610-A1.	
26-JUN-1997.	
20-DEC-1996;	96DE-1054610.
03-SEP-1996;	96US-0706945.
22-DEC-1995;	95US-0577788.
(AMGE-) AMGEN INC.	

XX Boyle WJ, Calzone FJ, Lacey DL, Chang MS;
XX WPI: 1997-334271/31.
XX N-PSDB: T96062.
XX Nucleic acid encoding osteoprotegerin - useful for treatment of
XX diseases involving excessive bone loss, e.g. osteoporosis
XX
XX Claim 23; Pages 106-107; 182pp; German.
XX
XX The present sequence is mouse osteoprotegerin (OPG). Anti-OPG
XX antibodies can be used in OPG diagnostic assays, and as affinity
XX purification materials. The OPG cDNA can be used to express
XX recombinant OPG and to generate transgenic animals. It can also
XX be used to regulate the level of OPG in mammals, specifically to
XX increase OPG levels, however the use of antisense sequences is
XX also contemplated. Fragments of the cDNA can be used as probes to
XX detect OPG expressing cells and tissue, and to screen cDNA
XX libraries for related sequences. OPG can be used to treat or
XX prevent bone diseases, specifically excessive bone loss, e.g.
XX osteoporosis, Paget's disease, hypercalcaemia,
XX hyperparathyroidism, rheumatoid arthritis, osteomyelitis, and
XX osteolytic metastases, periodontal bone loss, bone necrosis and
XX osteopaenia.
XX
XX Sequence 401 AA;
XX
XX Query Match 100.0%; Score 2179; DB 18; Length 401;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-179;
XX Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MNKWLCCALLVLLDIETWTQETLPKYLHYDPTGHLCDKCAPGYLKQHTVRKKT 60
XX Db 1 mnkwlccallvllldietwtqetlppkylhydpetghlclcdkcapgtylkqhtvrkt 60
XX
XX QY 61 LCVPDHSYTDSDWHTSDEVCYSPVKELQSVKQECNRTHNRVCEEGRYLEIEFCLK 120
XX Db 61 lcvpdhsytdsdwhtsdecyvspckelqsvkqecnrthnrviceeegryleiefclk 120
XX
XX QY 121 HRSPGPGSVVQAGTPERTVCKKCPDGFSGFTSSKAPCIKHTNCSTFGLLLLQKGNAT 180
XX Db 121 hrscppgpgsvvqagtpertvckkcpdgffsgftsskapcikhntcstfgllllqkgnat 180
XX
XX QY 181 HDNVCSGNREATQCGIDVTLCCEAFFRFAVPTKIIIPNWLVSVDLPGTKVNAESVERI 240
XX Db 181 hdnvcsgnreatqcgidvtlceaffrfavptkiiipnwlsvdlsipgtkvnaesveri 240
XX
XX QY 241 KRRHSSQEQTFOLLKWKHQRDQEMVKIIQDIDLCESSYORHLGHSNLTTEQLALME 300
XX Db 241 krrhssqeqtfollkwhqrdqemvkkiiqdidlcessygrhlghsnltteqlalme 300
XX
XX QY 301 SLPGKKISPEIERTRKTKSSBOLLKLLSWRIKNGDQDTLGLMYALKHLKTSHPFKT 360
XX Db 301 slpgkkspeiertrktsseqllkllslwrikngdqdtlglmyalkhlktshtpft 360
XX
XX QY 361 VTHSLRKTMRFLHSFTMYRLYKLFLEMGVQSVKTSCL 401
XX Db 361 vthslrktmrflhsftmyrlyqkflflemgvgvsktscl 401
XX
XX RESULT 2
XX W38343
XX ID W38343 standard; Protein; 401 AA.
XX AC W38343;
XX DT 20-APR-1998 (first entry)
XX DE Rat osteoprotegerin.
XX KW Rat; osteoprotegerin; antibody; diagnosis; affinity purification;

KW recombinant production; transgenic animal; treatment; prevention;
KW antisense oligonucleotide; probe; detection; screening;
KW bone disease; osteoporosis; Paget's disease; hypercalcaemia;
KW hyperparathyroidism; rheumatoid arthritis; osteomyelitis;
KW osteolytic metastasis; periodontal bone loss; bone necrosis;
KW osteopaenia.
XX Rattus sp.
XX OS
XX PN DE19654610-A1.
XX PD 26-JUN-1997.
XX PF 20-DEC-1996; 96DE-1054610.
XX PR 03-SEP-1996; 96US-0706945.
XX PR 22-DEC-1995; 95US-0577788.
XX (AMGE-) AMGEN INC.
XX Boyle WJ, Calzone FJ, Lacey DL, Chang MS;
XX WPI: 1997-334271/31.
XX N-PSDB: T96061.
XX Nucleic acid encoding osteoprotegerin - useful for treatment of
XX diseases involving excessive bone loss, e.g. osteoporosis
XX
XX Claim 23; Pages 102-104; 182pp; German.
XX
XX The present sequence is rat osteoprotegerin (OPG). Anti-OPG
XX antibodies can be used in OPG diagnostic assays, and as affinity
XX purification materials. The OPG cDNA can be used to express
XX recombinant OPG and to generate transgenic animals. It can also
XX be used to regulate the level of OPG in mammals, specifically to
XX increase OPG levels, however the use of antisense sequences is
XX also contemplated. Fragments of the cDNA can be used as probes to
XX detect OPG expressing cells and tissue, and to screen cDNA
XX libraries for related sequences. OPG can be used to treat or
XX prevent bone diseases, specifically excessive bone loss, e.g.
XX osteoporosis, Paget's disease, hypercalcaemia,
XX hyperparathyroidism, rheumatoid arthritis, osteomyelitis, and
XX osteolytic metastases, periodontal bone loss, bone necrosis and
XX osteopaenia.
XX
XX Sequence 401 AA;
XX
XX Query Match 95.4%; Score 2079; DB 18; Length 401;
XX Best Local Similarity 94.5%; Pred. No. 8.5e-171;
XX Matches 379; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
XX
XX QY 1 MNKWLCCALLVLLDIETWTQETLPKYLHYDPTGHLCDKCAPGYLKQHTVRKKT 60
XX Db 1 mnkwlccallvllldietwtqetlppkylhydpetghlclcdkcapgtylkqhtvrkt 60
XX
XX QY 61 LCVPDHSYTDSDWHTSDEVCYSPVKELQSVKQECNRTHNRVCEEGRYLEIEFCLK 120
XX Db 61 lcvpdhsytdsdwhtsdecyvspckelqsvkqecnrthnrviceeegryleiefclk 120
XX
XX QY 121 HRSPGPGSVVQAGTPERTVCKKCPDGFSGFTSSKAPCIKHTNCSTFGLLLLQKGNAT 180
XX Db 121 hrscppgpgsvvqagtpertvckkcpdgffsgftsskapcikhntcstfgllllqkgnat 180
XX
XX QY 181 HDNVCSGNREATQCGIDVTLCCEAFFRFAVPTKIIIPNWLVSVDLPGTKVNAESVERI 240
XX Db 181 hdnvcsgnreatqcgidvtlceaffrfavptkiiipnwlsvdlsipgtkvnaesveri 240
XX
XX QY 241 KRRHSSQEQTFOLLKWKHQRDQEMVKIIQDIDLCESSYORHLGHSNLTTEQLALME 300
XX Db 241 krrhssqeqtfollkwhqrdqemvkkiiqdidlcessygrhlghsnltteqlalme 300
XX
XX QY 301 SLPGKKISPEIERTRKTKSSBOLLKLLSWRIKNGDQDTLGLMYALKHLKTSHPFKT 360

Db 301 sipgkispdeiertrktckpseqllikllslwrkngdgtlkgmyalkhlykayhfpkt 360
 QY 361 vthslrktmrflhsftmyrlyoklflemignovqsvkiscsl 401
 Db 361 vthslrktmrflhsftmyrlyoklflemignovqsvkiscsl 401

RESULT 3

Y05742
 ID Y05742 standard; Protein; 401 AA.
 AC Y05742;
 DT 19-JUL-1999 (first entry)
 XX Tumour necrosis factor receptor TRL.
 DE Tumour necrosis factor receptor; TRL; osteoprotegerin; agonist;
 KW antagonist; screening; human; cancer; AIDS; Alzheimer's disease;
 KW inflammation; arthritis; septicemia; autoimmune disease;
 KW psoriasis; inflammatory bowel disease; transplant rejection;
 KW graft versus host disease; infection; stroke; ischemia;
 KW acute respiratory disease syndrome; restenosis; brain injury;
 KW bone disease; atherosclerosis; therapy.
 XX Homo sapiens.
 OS
 PN EP911633-A1.
 XX
 PD 28-APR-1999.
 PF 02-OCT-1998; 98EP-0203332.
 XX
 PR 08-OCT-1997; 97US-0061334.
 XX
 PA (SMK) SMITHKLINE BEECHAM CORP.
 XX
 PI McDonnell PC, Young PR, Zou J;
 XX
 DR WPI; 1999-246560/21.
 XX
 PT Identifying agonists and antagonists of tumor necrosis factor
 PT related receptors TRL, TR3 and TR5, and of ligand TL3, useful for
 PT treatment of cancer, AIDS, Alzheimer's disease, bone disease etc
 XX
 PS Disclosure; Page 10-12; 23pp; English.

XX The present sequence represents tumour necrosis factor receptor
 CC (TNFR) TRL, also known as osteoprotegerin. The invention relates
 CC to TNFR related polypeptides TRL, TR3 and TR5 (see Y05742-44) and
 CC their ligand TL3 (see Y05745). TRL, TR3, TR5 and TL3 are used in
 CC claimed methods of identifying agonists and antagonists, i.e.
 CC compounds that bind to the receptors or ligand, and which activate
 CC (agonist) or inhibit activation of (antagonists) TRL, TR3, TR5 or
 CC TL3. A screening kit for identifying agonists, antagonists,
 CC ligands, receptors, substrates, enzymes etc. for TRL, TR3, TR5 or
 CC TL3 polypeptides is provided. The agonists and antagonists are
 CC useful for treatment of chronic and acute inflammation, arthritis,
 CC septicemia, autoimmune disease e.g. inflammatory bowel disease,
 CC psoriasis, transplant rejection, graft versus host disease,
 CC infection, stroke, ischemia, acute respiratory disease syndrome,
 CC restenosis, brain injury, AIDS, bone diseases, cancer (e.g.
 CC lymphoproliferative disorders), atherosclerosis and Alzheimer's
 CC disease, etc., caused by imbalance of TRL, TR3, TR5 or TL3.

XX Sequence 401 AA;

Query Match 87.2%; Score 1900; DB 20; Length 401;
 Best Local Similarity 85.8%; Pred. No. 2.1e-155;
 Matches 345; Conservative 26; Mismatches 29; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLDI-I EWTTQETLPPKYLHYDPETGHQLLCKAPGYTLKQHCTVRRK 59
 Db 1 mnlkcca-lvfidisikwtgtfppkylhydeetshqllcdkcpptylkqhctakwk 59
 QY 60 TLCVPCPDHSYTSWHTSDECVYCSVKELQSVKQECNTHNRVCECEGRYLEIEFCL 119
 Db 60 tvcapcpdhyttdswhtsdeclycspvckelgyvkqecnrthrvceckeegyleiefcl 119
 QY 120 KHRCPPGGVVOAGTPERNTVCKKCPDGFSGETSSKAPCIKHTNCSTFGLLLIQKNA 179
 Db 120 khrcppggvgvqagtperntvckrcpdgffsnetsskpcrkhtncsvfllitqkna 179
 QY 180 THDNVCSGNREATQKCGIDVTLCCEAFFRPAVPTKIIPNWLSDVSLPOTKVNAESVER 239
 Db 180 thdnicsgnsestqcgidvtlceaffrfavptkftpnwslvldnlpotkvnaesver 239
 QY 240 IKRRHSSQEQTFOLLKLWKHQNDRQEMVKIIQDIDCESSVORHGLSHNLTEQLLALM 299
 Db 240 ikqrhssqeqtfqllklwkhnkqdkdivkiiqdidlccensvgrhghantfeqlrslm 299
 QY 300 ESTPGKKISPEETERTKCKSSPOLLLKLSLWRIKNGDODTLKGLMYALKHLKTSHPK 359
 Db 300 esipgkkgvgaediektackpsdqllkllslwrkngdodtlkglmhalkhsktyhpk 359
 QY 360 TVTHSLRKTRMRFLHSFTMYRLOKLFLEMIGNOVQSVKISCL 401
 Db 360 tvtqslkktirfihstmykykqlflemignovqsvkiscsl 401
 RESULT 4
 ID -W95030 standard; protein; 401 AA.
 AC W95030;
 XX
 DT 13-MAY-1999 (first entry)
 XX
 DE Tumour necrosis factor receptor (TNF-R) related polypeptide TRL.
 KW Tumour necrosis factor receptor; TNF-R; TRL; TR2; TL2; TL4; arthritis;
 KW inflammation; septicemia; autoimmune disease; transplant rejection;
 KW graft vs. host disease; infection; stroke; ischemia; brain injury; AIDS;
 KW acute respiratory disease syndrome; restenosis, bone disease; cancer;
 KW atherosclerosis; Alzheimer's disease.
 XX
 OS Unidentified.
 XX
 PN EP897114-A2.
 XX
 PD 17-FEB-1999.
 XX
 PF 04-JUN-1998; 98EP-0304424.
 XX
 PR 29-AUG-1997; 97US-0057550.
 PR 13-AUG-1997; 97US-0055513.
 PR 26-AUG-1997; 97US-0056980.
 XX
 PA (SMK) SMITHKLINE BEECHAM CORP.
 XX
 PI Brigham-burke MR, Young PR;
 XX WPI; 1999-134308/12.
 DR
 XX Identifying agonists and antagonists to tumour necrosis factor
 PT receptor (TNF-R) related polypeptides (TRL, TR2, TL2 and TL4) -
 PT useful for treating stroke, Alzheimer's disease and AIDS
 XX
 PS Disclosure; Page 11-12; 18pp; English.
 XX
 CC The invention relates to identifying agonists or antagonists to tumour
 CC necrosis factor receptor (TNF-R) related polypeptides (TRL and TR2, TL2
 CC and TL4). The method comprises: (a) (i) contacting TRL or TR2 with a
 CC candidate compound in the presence of TL2 or TL4; or (ii) contacting TL2

CC or TL4 with a candidate compound in the presence of TR1 or TR2; and (b)
 CC assessing the ability of the candidate compound to compete with TR1 or
 CC TR2 binding to TL4. TR and TL agonists and antagonists are useful
 CC for treating diseases caused by imbalance of TL or TR polypeptide levels,
 CC which cause: chronic and acute inflammation, arthritis, septicemia,
 CC autoimmune diseases, transplant rejection, graft vs. host disease,
 CC infection, stroke, ischemia, acute respiratory disease syndrome,
 CC restenosis, brain injury, AIDS, bone diseases, cancer, atherosclerosis
 CC and Alzheimer's disease. The present sequence represents a TNF-R related
 CC polypeptide TR1.
 XX
 SQ Sequence 401 AA;

Query Match 87.2%; Score 1900; DB 20; Length 401;
 Best Local Similarity 85.8%; Pred. No. 2.1e-155;
 Matches 345; Conservative 26; Mismatches 29; Indels 2; Gaps 2;
 QY 1 MNKWLCCALLVLDI-IEWTTOETLPPKYLHYDPETGHQLCDKCAPQTYLKQHCYVRRK 59
 DB 1 mnkllcca-lvldisikwtqetfppkylhydeetshqllcdkcpqgtylkqctakwk 59
 QY 60 TLCVPCPDHSDYTSWHTSDECVYSPVCKELQSVKQECNTHNRVCEEGRYLEIEFCL 119
 DB 60 tvcapcpdhyttdswhtsdeclycspvckelqyvkqecnrthnrvcceckegryleiefcl 119
 QY 120 KHRSCPPGSGVVOAGTPTERNVCKKCPDGPFSGETSSKAPCIKHTNCSTFGLLLIQKNA 179
 DB 120 khrcscppgvgvqagtperntvckrcpdgffsnetskapcrkhtncsvfllltqkqna 179
 QY 180 THDNVCSGNREATOKCGIDVTLCCEAFRRFAVPTKIIPNWLVLVDSLPQTKVNAESVER 239
 DB 180 thdnicsgnsesctqcgldvtlceeafrfavpctkftpnwlsvlvdlpqtgvnaesver 239
 QY 240 IKRRHSSQEQFTQLLKWKHONRQDQVVKTIQDIDICESVQRHLGHSNLTFTQLLALM 299
 DB 240 ikrrhssqeqftqllkwkhqndqdvkklqiddicensvqrhghantfeqlrslm 299
 QY 300 ESLPGKISPEIERTRKTKSSQQLKLLSLWRKNGDQDTLGLMVALKHLKTSHPFK 359
 DB 300 eslpgkvgaediektikackpsdqllkllslwrkngdqdtklglmhalkhsktyhfpk 359
 QY 360 TVTHSLRKTMRFLHSFTMYRLYQKLFLEMIGNQVQSVKISCL 401
 DB 360 tvtslktirflhsftmyrklyqlflemignqvsvkiscsl 401

RESULT 5
 W83926
 ID W83926 standard; Protein; 401 AA.
 AC AC
 AC W83926;
 XX
 DT 01-MAR-1999 (first entry)
 XX Human FTHMA-070 protein.
 DE Human FTHMA-070 protein.
 XX FTHMA-070; human; neurological disorder; therapy; diagnosis.
 KW Homo sapiens.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Sig_peptide
 FT Protein 22..401
 FT /label= Mat_protein
 XX
 PN W09848051-A2.
 XX
 PD 29-OCT-1998.
 XX
 PF 17-APR-1998; 98WO-US07714.
 XX

PR 10-OCT-1997; 97US-0062017.
 PR 18-APR-1997; 97US-0044746.
 XX
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
 XX
 PI Holtzman D, McCarthy SA;
 XX
 DR WPI; 1999-024021/02.
 DR N-PSDB; V69277.
 XX
 PT New isolated human FTHMA-070 and T85 proteins - used to develop
 PT products for the diagnosis and therapy of disorders involving
 PT cellular processes, e.g. neuronal development.
 XX
 PS Claim 8; Fig 1; 127pp; English.
 XX
 CC This is the amino acid sequence of human FTHMA-070, a novel protein
 CC having homology to tumour necrosis factor receptor. The sequence
 CC was deduced from that of a cDNA clone (see V69277) isolated from a
 CC cardiac coronary artery smooth muscle cell library. FTHMA-070
 CC nucleic acids and polypeptides of the invention are useful as
 CC modulating agents in regulating a variety of cellular processes.
 CC They can be used for identifying compounds which bind to or modulate
 CC the activity of the polypeptides (claimed). They can also be used
 CC in screening assays, detection assays (e.g. chromosomal mapping,
 CC tissue typing, forensic biology), predictive medicine (e.g.
 CC diagnostic assays, prognostic assays, monitoring clinical trials,
 CC and pharmacogenomics), and methods of treatment (e.g. therapeutic
 CC and prophylactic) e.g. for neurological disorders.
 XX
 SQ Sequence 401 AA;

Query Match 87.2%; Score 1900; DB 20; Length 401;
 Best Local Similarity 85.8%; Pred. No. 2.1e-155;
 Matches 345; Conservative 26; Mismatches 29; Indels 2; Gaps 2;
 QY 1 MNKWLCCALLVLDI-IEWTTOETLPPKYLHYDPETGHQLCDKCAPQTYLKQHCYVRRK 59
 DB 1 mnkllcca-lvldisikwtqetfppkylhydeetshqllcdkcpqgtylkqctakwk 59
 QY 60 TLCVPCPDHSDYTSWHTSDECVYSPVCKELQSVKQECNTHNRVCEEGRYLEIEFCL 119
 DB 60 tvcapcpdhyttdswhtsdeclycspvckelqyvkqecnrthnrvcceckegryleiefcl 119
 QY 120 KHRSCPPGSGVVOAGTPTERNVCKKCPDGPFSGETSSKAPCIKHTNCSTFGLLLIQKNA 179
 DB 120 khrcscppgvgvqagtperntvckrcpdgffsnetskapcrkhtncsvfllltqkqna 179
 QY 180 THDNVCSGNREATOKCGIDVTLCCEAFRRFAVPTKIIPNWLVLVDSLPQTKVNAESVER 239
 DB 180 thdnicsgnsesctqcgldvtlceeafrfavpctkftpnwlsvlvdlpqtgvnaesver 239
 QY 240 IKRRHSSQEQFTQLLKWKHONRQDQVVKTIQDIDICESVQRHLGHSNLTFTQLLALM 299
 DB 240 ikrrhssqeqftqllkwkhqndqdvkklqiddicensvqrhghantfeqlrslm 299
 QY 300 ESLPGKISPEIERTRKTKSSQQLKLLSLWRKNGDQDTLGLMVALKHLKTSHPFK 359
 DB 300 eslpgkvgaediektikackpsdqllkllslwrkngdqdtklglmhalkhsktyhfpk 359
 QY 360 TVTHSLRKTMRFLHSFTMYRLYQKLFLEMIGNQVQSVKISCL 401
 DB 360 tvtslktirflhsftmyrklyqlflemignqvsvkiscsl 401

RESULT 6
 R99925
 ID R99925 standard; Protein; 401 AA.
 XX
 AC R99925;
 XX
 DT 22-APR-1997 (first entry)

XX Full length osteoclastogenesis inhibitory factor.
 XX Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
 KW Osteoporosis.
 KW Homo sapiens.
 OS
 FH Key
 FT Peptide
 FT /note= "Signal peptide"
 FT Protein
 FT /note= "Mature OCIF, claim 6"
 FT
 PN WO9626217-A1.
 XX
 XX 29-AUG-1996.
 XX
 XX 20-FEB-1996; 96WO-JP00374.
 PF
 XX 21-JUL-1995; 95JP-0207508.
 PR
 PR 20-FEB-1995; 95JP-0054977.
 XX
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.
 XX
 XX Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
 XX
 XX WPI; 1996-402320/40.
 DR N-PSDB; T36685.
 XX
 XX DNA encoding osteoclastogenesis inhibitory factor protein - useful
 PT for bone resorption control, esp. treatment of osteoporosis
 PT
 XX Disclosure; Page 64-66; 183pp; Japanese.
 PS
 XX This sequence represents the full length osteoclastogenesis inhibitory
 CC factor (OCIF) of the invention. The OCIF has a molecular weight by
 CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-
 CC reducing conditions. The protein is adsorbed onto cation-exchangers
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or
 CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is
 CC useful in the control of bone resorption and therefore in the
 CC treatment and prevention of disorders of bone resorption, e.g.
 CC osteoporosis.
 CC
 XX Sequence 401 AA;
 SQ

Query Match 87.0%; Score 1895; DB 17; Length 401;
 Best Local Similarity 85.6%; Pred. No. 5.6e-155;
 Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;
 QY 1 MNKWLCCALLVLDI-TEWTTQETLPPKYLHYDPETGHLCDKCAPGTYLKQHTVRRK 59
 Db 1 mnnllcca-lvfidisikwtetfpkylhydeetshqllcdkcpptylkqhctakw 59
 QY 60 TLCVPCPDHSYTDSWHTSDSCVPCVKELQSVKQECNTHRVCEBGRYLEIEFCL 119
 Db 60 tvcaepcdpdytdswhtsdsclycspvckelqyvkqecnthrvceckegyleiefcl 119
 QY 120 KHRSCPPGSGVQAGTERTNVCKCPDGFSGTSSKAPCIKHTNCTFGLLLIQKGN 179
 Db 120 khrcpppgfvvgaqtpertntvckcpdgffsnetsskpcrkhtncsvfllltqkgn 179
 QY 180 THDNVCSGNREATQKCGIDVTLCEAEFFRAVPTKIIPNMLSLVDSLPCTKVNAESVER 239
 Db 180 thdnicsgsnestqcgldvtlceaeffravptkiiipnmlslvdslnpctkvnaesver 239
 QY 240 IKRHSSEQETFQLLKLWKNRQDMVKKLIQIDICSSVQRHLGHSNLTTEQLALM 299
 Db 240 ikrdhssgeqtfqllklwkhgnkdqdivkllqgidicssvqgrhlganltfeqlrslm 299

QY 300 ESLEPGKKISPEIETRTKCTCKSSQLKLKLSLWRKIKNGDQDTLKGLMYALKHLKTSHPK 359
 Db 300 eslpgkkvgaediektikackpsdqllklslwrkngdqtikglmhalkhsktyhfpk 359
 QY 360 TVVHSLRKTMRFILHSFTMYRLYOKLFLEMTGNQVQSVKISCL 401
 Db 360 tvtqslkktirflhsftmyklyqklflemignqvgvskiscl 401
 RESULT 7
 W53239
 ID W53239 standard; Protein; 401 AA.
 XX
 AC W53239;
 XX
 DT 15-JUL-1998 (first entry)
 XX
 DE Human OCIF genome DNA protein.
 XX
 KW Human; OCIF; genome; osteoclast; antipyretic; osteoporosis;
 KW rheumatism; multiple sclerosis.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /label= signal
 FT Protein
 FT /label= OCIF_protein
 FT
 XX WO9807840-A1.
 PN
 XX 26-FEB-1998.
 PD
 XX 19-AUG-1997; 97WO-JP02859.
 PF
 XX 19-AUG-1996; 96JP-0235928.
 PR (SNOW) SNOW BRAND MILK PROD CO LTD.
 PA
 XX Morinaga T, Nakagawa N, Yasuda H;
 PI WPI; 1998-169150/15.
 DR N-PSDB; V20768.
 DR
 XX Inhibition of osteoclast formation and/or antipyretic activity -
 PT useful for, e.g. treating osteoporosis, rheumatism and multiple
 PT sclerosis
 PT
 XX Claim 4; Page 26-28; 36pp; English.
 PS
 XX The present sequence represents human OCIF genome DNA protein which is
 CC specifically claimed in the present invention. The present invention
 CC describes: (1) a method of inhibiting the formation of osteoclasts
 CC and/or antipyretic proteins, which have the following characteristics:
 CC (i) MW determined by SDS-PAGE of approximately 60 kDa under reducing
 CC conditions, and 60 kDa and 120 kDa under non-reducing conditions; (ii)
 CC 401 aa sequence (III) (see W53239); (iii) affinity to cation exchangers
 CC and heparin; (iv) its osteoclast formation inhibiting activity reduced
 CC by heating at 70 degrees Celsius for 10 minutes or 56 degrees Celsius
 CC for 30 minutes, and (v) its antipyretic activity reduced by heating at
 CC 90 degrees Celsius for 10 minutes, and (2) a method for preparing the
 CC above proteins. The proteins are useful for, e.g. treatment and
 CC prevention of osteoporosis, rheumatism or multiple sclerosis, and also
 CC as antigens for immunological diagnosis of these diseases and disorders.
 XX
 SQ Sequence 401 AA;

Query Match 87.0%; Score 1895; DB 19; Length 401;
 Best Local Similarity 85.6%; Pred. No. 5.6e-155;
 Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

Qy 1 MNKWLCCALLVLLDI-IEWTTQETLPPKYLHYDPETHGQLLDCDKCAPGTYLKQHCTVRRK 59
 Db 1 mnnllcca-lvldisikwtqetfppkylhydeetsqllcdkcpptylkqhtakwk 59
 Qy 60 TLCVPCPDHSTYDSDWHTSDECVYSPVKELQSVKQECNRTHNRVCEEGRYLEIEFCL 119
 Db 60 tvacpcpdhytydswhtsdeclycspvckelqyvkqecnrthnrveckeegryleiefcl 119
 Qy 120 KHRSCPPGSGVVGQAGTPERNVCKKCPDGFSGTSSKAPCIKHTNCSFGILLIQKNA 179
 Db 120 khrcsppgfvvgagtperntvckrcpdgffsnetskapcrkhtncsvfgilltqkna 179
 Qy 180 THDNVCSGNREATOKCGIDVTLCCEAFPRFAVPTKIIPNWLVSILVDSLPQTKVNAESVER 239
 Db 180 thdnicsgnsestqkcgldvtlceaeaffrfaavtpktpnwslvldnlpqtkvnaesver 239
 Qy 240 IKRRHSQEQTFQLKLWKNQRDQEMVKIIQDIDLCSSVQRHLGHSNLTTEQLLALM 299
 Db 240 ikqhssqeqtfqlklwkhqkdqdvkviqddilcensvqrhghanitfeqlrslm 299
 Qy 300 ESLPGKKISPEIERTRKTCSSQOLKLLSLWRKNGDQDTLGLMYALKHLKTSHPFK 359
 Db 300 eslpqkkgvgaediektikackpsdqilklslwrkngdgtlkglmhalkhsktyhfpk 359
 Qy 360 TVTHSLRKTMRFLHSFTMYRYQLFLEMIGNOVQSVKISCL 401
 Db 360 tvtsqlkktirflhsftmyklyqkfllemignqvsvkiscsl 401

RESULT 8
 W38345
 ID W38345 standard; Protein; 401 AA.
 XX W38345;
 AC W38345;
 DT 20-APR-1998 (first entry)
 DE Human osteoprotegerin.
 XX Osteoprotegerin; antibody; diagnosis; affinity purification;
 KW recombinant production; transgenic animal; treatment; prevention;
 KW antisense oligonucleotide; probe; detection; screening; human;
 KW bone disease; osteoporosis; Paget's disease; hypercalcaemia;
 KW hyperparathyroidism; rheumatoid arthritis; osteomyelitis;
 KW osteolytic metastasis; periodontal bone loss; bone necrosis;
 KW osteopaenia.
 XX Homo sapiens.
 XX DE19654610-A1.
 XX 26-JUN-1997.
 XX 20-DEC-1996; 96DE-1054610.
 XX 03-SEP-1996; 96US-0706945.
 XX 22-DEC-1995; 95US-0577788.
 XX (AMGE-) AMGEN INC.
 XX Boyle WJ, Calzone FJ, Lacey DL, Chang MS;
 XX WPI; 1997-334271/31.
 XX N-FSDB; T96063.
 XX Nucleic acid encoding osteoprotegerin - useful for treatment of
 XX diseases involving excessive bone loss, e.g. osteoporosis
 XX Claim 23; Pages 109-111; 182pp; German.
 XX The present sequence is human osteoprotegerin (OPG). Anti-OPG
 XX antibodies can be used in OPG diagnostic assays, and as affinity
 XX purification materials. The OPG cDNA can be used to express

CC recombinant OPG and to generate transgenic animals. It can also
 CC be used to regulate the level of OPG in mammals, specifically to
 CC increase OPG levels, however the use of antisense sequences is
 CC also contemplated. Fragments of the cDNA can be used as probes to
 CC detect OPG expressing cells and tissue, and to screen cDNA
 CC libraries for related sequences. OPG can be used to treat or
 CC prevent bone diseases, specifically excessive bone loss, e.g.
 CC osteoporosis, Paget's disease, hypercalcaemia,
 CC hyperparathyroidism, rheumatoid arthritis, osteomyelitis,
 CC osteolytic metastases, periodontal bone loss, bone necrosis and
 CC osteopaenia.
 XX
 XX Sequence 401 AA;
 Qy
 Query Match 86.8%; Score 1892; DB 18; Length 401;
 Best Local Similarity 85.6%; Pred. No. le-154;
 Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;
 Qy 1 MNKWLCCALLVLLDI-IEWTTQETLPPKYLHYDPETHGQLLDCDKCAPGTYLKQHCTVRRK 59
 Db 1 mnnllcca-lvldisikwtqetfppkylhydeetsqllcdkcpptylkqhtakwk 59
 Qy 60 TLCVPCPDHSTYDSDWHTSDECVYSPVKELQSVKQECNRTHNRVCEEGRYLEIEFCL 119
 Db 60 tvacpcpdhytydswhtsdeclycspvckelqyvkqecnrthnrveckeegryleiefcl 119
 Qy 120 KHRSCPPGSGVVGQAGTPERNVCKKCPDGFSGTSSKAPCIKHTNCSFGILLIQKNA 179
 Db 120 khrcsppgfvvgagtperntvckrcpdgffsnetskapcrkhtncsvfgilltqkna 179
 Qy 180 THDNVCSGNREATOKCGIDVTLCCEAFPRFAVPTKIIPNWLVSILVDSLPQTKVNAESVER 239
 Db 180 thdnicsgnsestqkcgldvtlceaeaffrfaavtpktpnwslvldnlpqtkvnaesver 239
 Qy 240 IKRRHSQEQTFQLKLWKNQRDQEMVKIIQDIDLCSSVQRHLGHSNLTTEQLLALM 299
 Db 240 ikqhssqeqtfqlklwkhqkdqdvkviqddilcensvqrhghanitfeqlrslm 299
 Qy 300 ESLPGKKISPEIERTRKTCSSQOLKLLSLWRKNGDQDTLGLMYALKHLKTSHPFK 359
 Db 300 eslpqkkgvgaediektikackpsdqilklslwrkngdgtlkglmhalkhsktyhfpk 359
 Qy 360 TVTHSLRKTMRFLHSFTMYRYQLFLEMIGNOVQSVKISCL 401
 Db 360 tvtsqlkktirflhsftmyklyqkfllemignqvsvkiscsl 401
 RESULT 9
 Y43400
 ID Y43400 standard; Protein; 401 AA.
 XX Y43400;
 AC Y43400;
 DT 28-JAN-2000 (first entry)
 DE Osteoprotegerin protein sequence.
 XX Osteoprotegerin; OPG; human; cardiovascular disease; occlusion;
 KW calcification; blood vessel; atherosclerosis; medial calcific sclerosis;
 KW Monckeberg's arteriosclerosis; therapy.
 XX Homo sapiens.
 XX WO9953942-A1.
 XX 28-OCT-1999.
 XX 21-APR-1999; 99WO-US08793.
 XX 23-APR-1998; 98US-0064832.
 XX (AMGE-) AMGEN INC.
 PA

XX Simonet S, Sarosi I;
 XX WPI: 2000-013182/01.
 DR N-PSDB; 237254.
 XX
 PT Treating and preventing cardiovascular diseases, especially
 PT atherosclerosis and Monckeberg's arteriosclerosis
 XX
 XX Claim 9; Page 37-39; 43pp; English.
 XX
 CC This sequence represents the human osteoprotegerin (OPG). The invention
 CC relates to a method of treating or preventing cardiovascular disease by
 CC administering OPG. The method can be used to treat and prevent
 CC cardiovascular diseases associated with occlusion and calcification of
 CC blood vessels, especially atherosclerosis or Monckeberg's
 CC arteriosclerosis, i.e. medial calcific sclerosis. Using OPG to treat or
 CC prevent cardiovascular diseases provides an alternative to invasive
 CC treatments. OPG can be used as a single therapeutic for prevention and
 CC treatment of both osteoporosis and cardiovascular diseases.
 XX
 XX Sequence 401 AA;
 SQ
 Query Match 86.8%; Score 1892; DB 21; Length 401;
 Best Local Similarity 85.3%; Pred. No. 1e-154;
 Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;
 QY 1 MNKWLCCALLVLLDI-IEWTTQETLPPKYLHYDPETHGQLLCKAPGYLYKHQCHVRRK 59
 DB 1 mnkllcca-lvfidisikwtgtfppkylhydeetshgllcdkcpptgylkqhctakwk 59
 QY 60 TLVCPCHSYTDSWHTSDCYVSPVCKELQSVKQECNTHNRVCECEGRVLEIEFCL 119
 DB 60 tvcapcdpdyttdswhtsdclcyvckelgyvkcenrthnrvcceckegrylelefc 119
 QY 120 KHRSCPPGGVQAGTPERTVCKKCPDGFSGTSSKAPCIKHTNCSFTGLLLIQKNA 179
 DB 120 khrcppggvqvagtpertntvckrcpdgffsnetsskapckrhtncsvfgllltqkna 179
 QY 180 THNVCSGNREATQKCGIDVTLCCEAFFRFAVPTKIIPNWLVLDSLPOTKVNASVER 239
 DB 180 thnicsgnsctqkcgidvtlceeafrfavptkftpnwslvldnlpotkvnaesver 239
 QY 240 IKRRHSSQEQTFOLLKWLKHONRQEMVKIIQDIDLCSSVORHGLHNSLTTEQLLALM 299
 DB 240 ikrqhssqeqtfllkwlkhqnaqdvkvkllqdidlcensvgrhghnaltfeqlrslm 299
 QY 300 ESLPGKKISPEIETRTKCKTSSEQLLKLSLWRIRKNGDQDTLGLMYALKHLKTSHPFK 359
 DB 300 eslpgkkvgaediektikackpsdqllkllslwrknqdgtikglmhalkhsktyhfpk 359
 QY 360 TVTHSLRKTMRFLHSFTMYLYOKLFLEMIGNQVQSKISCL 401
 DB 360 tvtqslkktirflhsftmylyqkflflemignvqvskiscl 401
 RESULT 10
 W57635
 ID W57635 standard; Protein; 401 AA.
 XX
 AC W57635;
 XX
 DT 27-AUG-1998 (first entry)
 XX
 DE Trl receptor protein.
 XX
 KW Trl receptor; tumour necrosis factor receptor-1; TNF-alpha; TNF-beta;
 KW inhibitor; tumour growth; tumour necrosis; microorganism infection;
 KW cellular differentiation stimulation; ionising radiation; septic shock;
 KW anti-viral response; growth regulator; immune response; meningococemia;
 KW autoimmune disease; apoptosis; inflammation; cachexia; cerebral malaria;
 KW AIDS; therapy.

XX Homo sapiens.
 XX WO9812344-A1.
 XX
 XX 26-MAR-1998.
 XX
 XX 18-SEP-1996; 96WO-US15003.
 XX
 XX 18-SEP-1996; 96WO-US15003.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Fleischmann RD, Greene JM, Ni J;
 PI
 XX WPI: 1998-217278/19.
 DR N-PSDB; V24486.
 XX
 XX New isolated tumour necrosis factor receptor - useful for developing
 XX products for treating, e.g. tumours, auto-immune disease(s), graft
 XX rejection, apoptosis or inflammation
 XX
 XX Claim 13; Fig 1; 11pp; English.
 XX
 CC This sequence is a tumour necrosis factor receptor-1 (Trl) receptor
 CC of the invention. The DNA can be used to produce a recombinant host
 CC cell by inserting it into a vector, which is then used to transfect the
 CC host cell. The Trl receptor can bind both TNF-alpha and TNF-beta. Trl
 CC receptor agonists can be used for inhibition of tumour growth and
 CC necrosis of tumours. They can also be used to stimulate cellular
 CC differentiation, e.g. T cell, fibroblasts or haematopoietic cell
 CC differentiation, may be used to augment Trl's role in the host's defence
 CC against microorganisms and prevent related disease. The agonists may
 CC also be used to protect against the deleterious effects of ionising
 CC radiation produced during a course of radiotherapy, e.g. denaturation of
 CC enzymes, lipid peroxidation or DNA damage. The agonists may further be
 CC used to mediate an anti-viral response, to regulate growth, to mediate
 CC the immune response and to treat immunodeficiencies related to diseases
 CC such as HIV. Antagonists to the Trl receptor may be used to treat
 CC autoimmune diseases, e.g. graft versus host rejection and allograft
 CC rejection, and T cell mediated autoimmune diseases. They may also be used
 CC to prevent apoptosis. They may also be used to prevent cytotoxicity and
 CC to treat septic shock, meningococemia, inflammation, bacterial
 CC infections, cachexia, cerebral malaria or AIDS. The products can also be
 CC used for diagnosing the above diseases.
 XX
 XX Sequence 401 AA;
 SQ
 Query Match 86.8%; Score 1891; DB 19; Length 401;
 Best Local Similarity 85.3%; Pred. No. 1.2e-154;
 Matches 343; Conservative 27; Mismatches 30; Indels 2; Gaps 2;
 QY 1 MNKWLCCALLVLLDI-IEWTTQETLPPKYLHYDPETHGQLLCKAPGYLYKHQCHVRRK 59
 DB 1 mnkllcca-lvfidisikwtgtfppkylhydeetshgllcdkcpptgylkqhctakwk 59
 QY 60 TLVCPCHSYTDSWHTSDCYVSPVCKELQSVKQECNTHNRVCECEGRVLEIEFCL 119
 DB 60 tvcapcdpdyttdswhtsdclcyvckelgyvkcenrthnrvcceckegrylelefc 119
 QY 120 KHRSCPPGGVQAGTPERTVCKKCPDGFSGTSSKAPCIKHTNCSFTGLLLIQKNA 179
 DB 120 khrcppggvqvagtpertntvckrcpdgffsnetsskapckrhtncsvfgllltqkna 179
 QY 180 THNVCSGNREATQKCGIDVTLCCEAFFRFAVPTKIIPNWLVLDSLPOTKVNASVER 239
 DB 180 thenicsgnsctqkcgidvtlceeafrfavptkftpnwslvldnlpotkvnaesver 239
 QY 240 IKRRHSSQEQTFOLLKWLKHONRQEMVKIIQDIDLCSSVORHGLHNSLTTEQLLALM 299
 DB 240 ikrqhssqeqtfllkwlkhqnaqdvkvkllqdidlcensvgrhghnaltfeqlrslm 299

Qy 300 ESLPGKKISPEIERTRKTKSSQLLKLKLSLRIRKNGDQDTLGLMYALKHLKTSHPK 359
 Db 300 eslpqkvgaediektkackpsdqllklsrlwrikngdqtllgmlhalkhsktyhfpk 359
 Qy 360 TVTSLRKTMRFLHSFTMYRKYQLFLMIGNOVQSVKISCL 401
 Db 360 tvtdslkktirflhsftmyklyqlflmignqvsvkisc 401

RESULT 11
 ID R99932 standard; Protein; 401 AA.
 AC R99932;
 DT 22-APR-1997 (first entry)
 DE Mutated OCIF, OCIF-C20S.
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
 KW osteoporosis.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21 /note= "Signal peptide"
 FT Protein 22..401 /note= "Mature OCIF-C20S"
 FT Misc-difference 202 /label= C20S
 FT
 XX
 PN WO9626217-A1.
 XX
 PD 29-AUG-1996.
 XX
 PF 20-FEB-1996; 96WO-JP00374.
 XX
 PR 21-JUL-1995; 95JP-0207508.
 PR 20-FEB-1995; 95JP-0054977.
 XX
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.
 XX
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
 XX
 DR WPI; 1996-402320/40.
 DR N-PSDB; T33162.
 XX
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
 PT for bone resorption control, esp. treatment of osteoporosis
 XX
 PS Claim 32; Page 96-98; 183pp; Japanese.
 XX
 CC This sequence represents a mutated version of the full length
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
 CC sequence represents OCIF-C20S in which the 20th Cys residue in the
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
 CC cation-exchangers or heparin and its activity is lowered after 10 mins
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
 CC deg.C. OCIF is useful in the control of bone resorption and therefore
 CC in the treatment and prevention of disorders of bone resorption, e.g.
 CC osteoporosis.
 XX
 SQ Sequence 401 AA;

Query Match 86.5%; Score 1885; DB 17; Length 401;
 Best Local Similarity 85.3%; Pred. No. 4.1e-154;
 Matches 343; Conservative 26; Mismatches 31; Indels 2; Gaps 2;

Qy 1 MNKWLCCALLVLDI-IEWTQETLPPKYLHYDPETGHLQCLDKCAPGTYLKQHCTVRRK 59
 Db 1 mnllcca-lvldisikwtqetfppkylhydeetsqlldckcpptylkqhtakwk 59
 Qy 60 TLCVPCPDHSYTDSWHTSDCYCSPCKELQSVKQECNRTHNRVCECEEGRYLEIEFCL 119
 Db 60 tvcapcpdhytytdswhtsdeclycspvckelqyvkqecnrthnrvceckegryleiefcl 119
 Qy 120 KHRSCPPGSGVVOAGTPTERTVCKKCPDGFSGETSSKAPCIKHTNCSTFGILLIQGNA 179
 Db 120 khrcpppgfvgvqgtperntvckrcpdgffsnetsskpcrkhnscsvfgillcqgna 179
 Qy 180 THDNVCSGNREATQCGIDVTLCCEAFRRFAVPTKIIPNWLISVLVDSLPFGKVNAAESYER 239
 Db 180 thdnicsgnsestqcgidvtlseeaffravptkftpnwlsvldnlpqtkvnaesver 239
 Qy 240 IKRRHSQEQTFQLLKLWKHONROEMVKYIQDIDLCSSVQRHLGHSNLTTEQLLALM 299
 Db 240 ikrghsseqtfqlklwkhnkdqdvkikiqdidlcensvqrhghantfeqlrslm 299
 Qy 300 ESLPGKKISPEIERTRKTKSSQLLKLKLSLRIRKNGDQDTLGLMYALKHLKTSHPK 359
 Db 300 eslpqkvgaediektkackpsdqllklsrlwrikngdqtllgmlhalkhsktyhfpk 359
 Qy 360 TVTSLRKTMRFLHSFTMYRKYQLFLMIGNOVQSVKISCL 401
 Db 360 tvtdslkktirflhsftmyklyqlflmignqvsvkisc 401

RESULT 12
 ID R99931 standard; Protein; 401 AA.
 AC R99931;
 DT 22-APR-1997 (first entry)
 DE Mutated OCIF, OCIF-C19S.
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
 KW osteoporosis.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21 /note= "Signal peptide"
 FT Protein 22..401 /note= "Mature OCIF-C19S"
 FT Misc-difference 195 /label= C19S
 FT
 XX
 PN WO9626217-A1.
 XX
 PD 29-AUG-1996.
 XX
 PF 20-FEB-1996; 96WO-JP00374.
 XX
 PR 21-JUL-1995; 95JP-0207508.
 PR 20-FEB-1995; 95JP-0054977.
 XX
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.
 XX
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
 XX
 DR WPI; 1996-402320/40.
 DR N-PSDB; T33161.
 XX
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
 PT for bone resorption control, esp. treatment of osteoporosis
 XX
 PS Claim 29; Page 94-96; 183pp; Japanese.

XX This sequence represents a mutated version of the full length
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
 CC sequence represents OCIF-CL95 in which the 19th Cys residue in the
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
 CC cation-exchangers or heparin and its activity is lowered after 10 mins
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
 CC deg.C. OCIF is useful in the control of bone resorption and therefore
 CC in the treatment and prevention of disorders of bone resorption, e.g.
 CC osteoporosis.

XX Sequence 401 AA;

Query Match 86.5%; Score 1885; DB 17; Length 401;
 Best Local Similarity 85.3%; Pred. No. 4.1e-154;
 Matches 343; Conservative 26; Mismatches 31; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLLDI-IEWTTQETLPKYLHYDPETGHQLLCKCAPGTYLKQHCTVRRK 59
 DB 1 mnnllcca-lvldisikwtqetfppkylhydeetshqldckppgtylkqhtakwk 59
 QY 60 TLCVPCPDHSDYSDSWHTSDECVGSPVCKELQSVKQECNTHNRVCEEGRYLEIEFCL 119
 DB 60 tvcapcpdhytswhtsdeclcsyckelqyvkqecnrthnrvccekegryleiefcl 119
 QY 120 KHRSCPPGSGVQAGTPERNVCKKCPDGFSGTSSKAPCIKHTNCSTFGLLLIQKNA 179
 DB 120 khrrscppgsgvqagtperntvckrcpdgffsnetsskapckrhtncsvfllltqkna 179
 QY 180 THDNVCSGNREATQCGIDVTLCBEAFRAFPVKIIPNWLVSVDLSLPGTKVNAESVER 239
 DB 180 thdnicsgnsstqsgldvtlceaeaffravpckfnpwlsvldnlpgtkvnaesver 239
 QY 240 IKRRHSSQEQTFOLLKWKHQRNDQEMVKKIOTDIDCESSVQRHLGHSNLTTEQLLALM 299
 DB 240 ikrrhssqeqtfqllkwwkqndqdivkikiqdidcensvqrhighanltfeqlslm 299
 QY 300 ESLPGKKISPEIERTRKTKSSQQLKLLSLWRKNGDQDTLGLMYALKHLKTSHPFK 359
 DB 300 eslpqkkgvgaediektikackpsdqilkslwrlkngdgtlkgmlhalkhsktyhfpk 359
 QY 360 TVTSLRKTMRFLHSFTMYRLYQKLFLEMIGNOVQSVKISCL 401
 DB 360 tvtslrlktirflhsftmyrlyqkflflemignovqsvkiscl 401

RESULT 13
 R99942
 ID R99942 standard; Protein; 399 AA.
 XX AC R99942;
 XX 23-APR-1997 (first entry)
 XX Mutated OCIF, OCIF-CL.
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
 KW osteoporosis.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Peptide 1..21
 FT /note= "Signal peptide"
 FT Protein 22..399
 FT /note= "Mature OCIF-CL"
 XX W09626217-A1.
 XX 29-AUG-1996.

XX 20-FEB-1996; 96WO-JP00374.
 XX 21-JUL-1995; 95JP-0207508.
 PR 20-FEB-1995; 95JP-0054977.
 XX (SNOW) SNOW BRAND MILK PROD CO LTD.
 XX Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
 XX WPI; 1996-402320/40.
 DR N-PSDE; T33172.
 XX DNA encoding osteoclastogenesis inhibitory factor protein - useful
 PT for bone resorption control, esp. treatment of osteoporosis
 PT Claim 62; Page 117-119; 183pp; Japanese.
 PS This sequence represents a mutated version of the full length
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
 CC sequence represents OCIF-CL in which amino acids 379-380 of the
 CC mature OCIF protein are deleted. The OCIF of the invention
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
 CC cation-exchangers or heparin and its activity is lowered after 10 mins
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
 CC deg.C. OCIF is useful in the control of bone resorption and therefore
 CC in the treatment and prevention of disorders of bone resorption, e.g.
 CC osteoporosis.
 XX Sequence 399 AA;

Query Match 86.4%; Score 1882; DB 17; Length 399;
 Best Local Similarity 85.5%; Pred. No. 7.3e-154;
 Matches 342; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLLDI-IEWTTQETLPKYLHYDPETGHQLLCKCAPGTYLKQHCTVRRK 59
 DB 1 mnnllcca-lvldisikwtqetfppkylhydeetshqldckppgtylkqhtakwk 59
 QY 60 TLCVPCPDHSDYSDSWHTSDECVGSPVCKELQSVKQECNTHNRVCEEGRYLEIEFCL 119
 DB 60 tvcapcpdhytswhtsdeclcsyckelqyvkqecnrthnrvccekegryleiefcl 119
 QY 120 KHRSCPPGSGVQAGTPERNVCKKCPDGFSGTSSKAPCIKHTNCSTFGLLLIQKNA 179
 DB 120 khrrscppgsgvqagtperntvckrcpdgffsnetsskapckrhtncsvfllltqkna 179
 QY 180 THDNVCSGNREATQCGIDVTLCBEAFRAFPVKIIPNWLVSVDLSLPGTKVNAESVER 239
 DB 180 thdnicsgnsstqsgldvtlceaeaffravpckfnpwlsvldnlpgtkvnaesver 239
 QY 240 IKRRHSSQEQTFOLLKWKHQRNDQEMVKKIOTDIDCESSVQRHLGHSNLTTEQLLALM 299
 DB 240 ikrrhssqeqtfqllkwwkqndqdivkikiqdidcensvqrhighanltfeqlslm 299
 QY 300 ESLPGKKISPEIERTRKTKSSQQLKLLSLWRKNGDQDTLGLMYALKHLKTSHPFK 359
 DB 300 eslpqkkgvgaediektikackpsdqilkslwrlkngdgtlkgmlhalkhsktyhfpk 359
 QY 360 TVTSLRKTMRFLHSFTMYRLYQKLFLEMIGNOVQSVKIS 399
 DB 360 tvtslrlktirflhsftmyrlyqkflflemignovqsvkis 399

RESULT 14
 R99934
 ID R99934 standard; Protein; 401 AA.
 XX AC R99934;
 XX

DT 22-APR-1997 (first entry)
XX Mutated OCIF, OCIF-C22S.
DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
XX Synthetic.
OS Key Location/Qualifiers
FH Peptide 1..21 /note= "Signal peptide"
FT Protein 22..401 /note= "Mature OCIF-C22S"
FT Misc-difference 277 /label= C22S
FT /label= C22S
XX WO9626217-A1.
PN 29-AUG-1996.
XX 20-FEB-1996; 96WO-JP00374.
XX 21-JUL-1995; 95JP-0207508.
XX 20-FEB-1995; 95JP-0054977.
XX (SNOW) SNOW BRAND MILK PROD CO LTD.
XX Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
XX WPI; 1996-402320/40.
XX N-PSDB; T33164.
XX DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
XX Claim 38; Page 100-102; 183pp; Japanese.
XX This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C22S in which the 22nd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
XX osteoporosis.
XX Sequence 401 AA;
Query Match 86.3%; Score 1880; DB 17; Length 401;
Best Local Similarity 85.1%; Pred. No. 1.le-153;
Matches 342; Conservative 26; Mismatches 32; Indels 2; Gaps 2;
QY 1 MNKWLCCALLVLLDI-IEWTQETLPPKLYHDPTGTHQLLDCAPGTLYLKQCHTVRRK 59
DB 1 mnnllcca-lvldisikwtgtfpkylhydeetshgllcdkcpptgylkghctakwk 59
QY 60 TLCVPCPDHSTYDTSWHTDECVCYSPVCKELQSVKQECNRTHNRVCEEGRYLEIEFCL 119
DB 60 tvacpdpdhytdswhtsdclcyvspckelqykvqecnrthnrvccekgryleiefcl 119
QY 120 KHRSCPPGSGVQAGTPERNTVCKKCPDGFSGFTSSKAPCIKHTNCSTFGLLLLQKQNA 179
DB 120 khrcpppgfvvgaqtperntvckrcpdgffsnetsskpcrkhtncsvfllllqkgqna 179
QY 180 THDNCVCSNREATQCGIDVLCIEAFRAFPFKLIENWLSVLVDLPGTKVNAESYER 239
DB 180 thdnicsgnsstqkcgidvltlceaeaffravptkfnwlsvlvdlpgtkvnaesver 239

QY 240 IKRRHSSQETFQLLKLWKHQNRDQEMVKKIIQDIDLCSSVQRHGLHSNLTTEQLALM 299
DB 240 lkrhssqeqtfllklwkhnkqkddvkkiiqddlcensvqrhghlanltfeqlrslm 299
QY 300 ESLPGKKISPEIERTRKTSKSSQELLKLLSLWRKNGDQDTLGLMYALKHLKTSHPFK 359
DB 300 eslpgkkgvgaediektikaskpsdqilnllslwrkngdqdtlglmlhalkhsktyhfpk 359
QY 360 TVTHSLRKTMRPLHSFTMYRLYQKLFLEMGNOVSQSVKISCL 401
DB 360 tvtqslkktirflhsftmyklyqkflflemignqgsvskiscl 401
RESULT 15
R99933 ID R99933 standard; Protein; 401 AA.
XX AC R99933;
XX 22-APR-1997 (first entry)
DT Mutated OCIF, OCIF-C21S.
DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
XX KW osteoporosis.
XX OS Synthetic.
XX Key Location/Qualifiers
FH Peptide 1..21 /note= "Signal peptide"
FT Protein 22..401 /note= "Mature OCIF-C21S"
FT Misc-difference 277 /label= C21S
XX WO9626217-A1.
XX 29-AUG-1996.
XX 20-FEB-1996; 96WO-JP00374.
XX 21-JUL-1995; 95JP-0207508.
XX 20-FEB-1995; 95JP-0054977.
XX (SNOW) SNOW BRAND MILK PROD CO LTD.
XX Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
XX WPI; 1996-402320/40.
XX N-PSDB; T33163.
XX DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
XX Claim 35; Page 98-100; 183pp; Japanese.
XX This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C21S in which the 21st Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
XX osteoporosis.
XX Sequence 401 AA;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:13:22 ; Search time 135.68 Seconds
(without alignments)
200.414 Million cell updates/sec

Title: US-09-389-782A-1

Perfect score: 1263

Sequence: 1 EPKCDKTHTCPPCAPPELL.....MHEALHNYTKQSLSPGK 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_15:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_protist:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	829.5	65.7	437	11 Q9RIA4	Q9RIA4 mus musculus
2	272.5	21.6	684	13 Q90544	Q90544 ginglymosto
3	272.5	21.5	384	4 Q9UP60	Q9UP60 homo sapien
4	268	21.2	416	4 Q9NPP6	Q9NPP6 homo sapien
5	183	14.5	509	11 Q9GX57	Q9GX57 mus musculus
6	183	14.5	513	11 P97797	P97797 mus musculus
7	181	14.3	268	13 Q90524	Q90524 ginglymosto
8	179	14.2	506	6 O46632	O46632 bos taurus
9	177	14.0	509	11 O08907	O08907 mus musculus
10	175	13.9	506	6 O46631	O46631 bos taurus
11	171.5	13.6	259	13 Q90530	Q90530 ginglymosto
12	171.5	13.6	509	11 Q9WTN4	Q9WTN4 mus musculus
13	169	13.4	238	7 Q9MXA2	Q9MXA2 aulonocara
14	166.5	13.2	257	13 Q90536	Q90536 ginglymosto
15	165	13.1	237	13 Q90545	Q90545 ginglymosto
16	162.5	12.9	252	13 Q90568	Q90568 ginglymosto
17	162.5	12.9	267	13 Q90529	Q90529 ginglymosto
18	161	12.7	261	7 O19363	O19363 sus scrofa
19	156.5	12.4	237	7 Q9MX99	Q9MX99 aulonocara

20	154	12.2	208	7 Q9MXA0	Q9MXA0 aulonocara
21	154	12.2	261	7 O62868	O62868 sus scrofa
22	146.5	11.6	354	4 Q9NOK8	Q9NOK8 homo sapien
23	145.5	11.5	261	7 O98263	O98263 sus scrofa
24	144	11.4	397	4 Q9Y4V0	Q9Y4V0 homo sapien
25	143.5	11.4	248	7 O98044	O98044 aulonocara
26	143.5	11.4	354	4 Q9PLW8	Q9PLW8 homo sapien
27	143	11.3	214	11 Q9RIA5	Q9RIA5 mus musculus
28	143	11.3	260	7 P79551	P79551 homo sapien
29	142.5	11.3	183	7 Q29922	Q29922 homo sapien
30	142.5	11.3	261	7 Q30099	Q30099 homo sapien
31	142	11.2	183	7 Q30102	Q30102 homo sapien
32	142	11.2	398	4 O00241	O00241 homo sapien
33	141.5	11.2	183	7 Q30090	Q30090 homo sapien
34	141.5	11.2	229	7 Q30075	Q30075 homo sapien
35	141.5	11.2	238	7 Q31467	Q31467 morone saxa
36	141	11.2	237	7 Q31468	Q31468 morone saxa
37	141	11.2	237	7 Q31470	Q31470 morone saxa
38	141	11.2	261	7 Q30091	Q30091 homo sapien
39	140.5	11.1	188	7 Q95466	Q95466 homo sapien
40	140	11.1	229	7 Q30077	Q30077 homo sapien
41	140	11.1	229	7 Q30078	Q30078 homo sapien
42	140	11.1	237	7 Q31469	Q31469 morone saxa
43	140	11.1	261	7 Q31633	Q31633 homo sapien
44	140	11.1	264	7 Q29826	Q29826 homo sapien
45	140	11.1	273	7 Q29746	Q29746 homo sapien

ALIGNMENTS

```
RESULT 1
Q9RIA4 PRELIMINARY; PRT; 437 AA.
ID Q9RIA4;
AC Q9RIA4;
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-OCT-2000 (TREMREL. 15, Last annotation update)
DE GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1;
DR HSSP; P01842; 7FAB.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; ig; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;
```

```
Query Match 65.7%; Score 829.5; DB 11; Length 437;
Best Local Similarity 61.4%; Pred. No. 5.7e-68;
Matches 143; Conservative 44; Mismatches 37; Indels 9; Gaps 3;

QY 2 PKSCDKTHTCPP--CPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDSHEDPEVK 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 PRDGG-----CKPCICTVPEV---SSVFPPKPKDVLITLTPKVTCCVVDDISKDDPEVQ 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 FNVYDGVGVHNAKTRPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEK 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 265 FSWFDDVVEYHTAQTPREQFNSTFRSFLPTMHODWLNKGEYKCKVSNKALPAPIEK 324
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 TISKAKQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 179
```

Db 325 TISKTKGRKAPQVYTIPTPPKQMAKDVSLTCTMTDFPDTITVWQWNGQPAENYKNT 384

Qy 180 PVLDSGDSFFLYSKLTVDKSRWQGVFSCVSMHEALHNHYTKSLSPCK 232

Db 385 QPIMTDGSIYFISKLNVKQSWNEAGNTFTCSVLHGLHNHTEKLSHSPCK 437

RESULT 2

Q90544 PRELIMINARY; PRT; 684 AA.
AC Q90544;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.
OS Ginglymostoma cirratum (Nurse Shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RX MEDLINE-95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks."
RL Nature 374:168-173(1995).
DR EMBL; U18701; AAB48195.1; -.
DR HSSP; P01857; 1FC1.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; ig; 6.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DEF CRC64;

Query Match 21.6%; Score 272.5; DB 13; Length 684;
Best Local Similarity 33.3%; Pred. No. 9.4e-17;
Matches 71; Conservative 35; Mismatches 98; Indels 9; Gaps 7;

Qy 24 SVFLPPKPKDTLMSRPEVTCVVVDVSHEDPE-VKENWYVDGVEVHNATKPREQYN 82

Db 459 SVSLKLP-PFBEITWQQTATVCEIV---YSDLENIKVFWQVNGVERKKGVETQNPWSG 514

Qy 83 STYRVSVLTVLHODWLNKGYKCVSNKALPAIETKTSKAK-GQPREPQVYTLPPSRD 141

Db 515 SKSTIVSKLVYASWDSGTEVCLVDESELTPTPKASIRKANSQMHPPKVVLLHPSTD 574

Qy 142 EL-TKNQVSLCLVKGFYPSDIAEWESNGO-PENNYKTTTPVLDSGDSFFLYSKLTVDK 199

Db 575 EIDTENSATLMCLATNFHFAEIVYGVGMANDTLDSGYRTQVDSEKSGGSFVTDRLUTA 634

Qy 200 SRWQGNVFCSCVMHEALHNHYTKSLSPCK 232

Db 635 AEWNSDITTSCLVGHPSL-NRDLINKSTNKSNGK 666

RESULT 3

Q9UP60 PRELIMINARY; PRT; 384 AA.
AC Q9UP60;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE SNC73 PROTEIN.
GN SNC73.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;
RT "Identification and characterization of SNC73, a gene which is down-
RT regulated in colorectal cancer."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067420; AAC19365.1; -.
DR HSSP; P01825; 7FAB.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;

Query Match 21.5%; Score 272; DB 4; Length 384;
Best Local Similarity 31.2%; Pred. No. 5.1e-17;
Matches 79; Conservative 33; Mismatches 109; Indels 32; Gaps 10;

Qy 7 KTHTCP-----PCPAPELLGSPVLEFPKPK-----DTLMISKTPEVT 45

Db 119 RHYTNPSQDVTVPCEVPTPTTSPST-PPTPSPSCCHRLSLHRLPALEDLILLSEANLT 177

Qy 46 CVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGYK 105

Db 178 CTLTGL-RDASGVFTTWPSSGK--SAVQGPPELDLCYSSVSLPCGAEPWNGHKTFT 234

Qy 106 CKVSNKALPAIETKTSKAKGPREPQVYTLPPSRDELTKNO-VSLTCLVKGFYPSDIAV 164

Db 235 CTAAYPESTKTLTATLSKS-GNTERPEVHLLPPPPSEELALNELVTLTCLARGFSKDLV 293

Qy 165 EWESNGO--PENNYKTTTPVLVD-SDG--SEFLYSLKLTVDKSRWQGNVFCSCVMHEALHN 219

Db 294 RWLQSQELPREKYLTVASROEPSQGTTFVAVTSLRVAEDWKKGDTFSCWVGHEALPL 353

Qy 220 HYTKSLSPCK 232

Db 354 AFTQKTDRLAGK 366

RESULT 4

Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansoorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RT "The European IMAGE consortium for integrated Molecular analysis of
RT human gene transcripts."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389978; CAB97534.1; -.
FT NON_TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8CA1708BB8AB4687 CRC64;

Query Match 21.2%; Score 268; DB 4; Length 416;
Best Local Similarity 32.8%; Pred. No. 1.3e-16;
Matches 77; Conservative 34; Mismatches 112; Indels 12; Gaps 9;

```

QY 4 SCDKTHCPCPAPPELLGGSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWY 63
Db 170 SODVTVPVCPVPPPPCC-HPRLSLHRPALED-LLLGSEANLTLTGL-RDASGATFTWT 226

QY 64 VDGVEVHNKTKPREQYNSTYRVSVLTVLHODWLNKREYKCKVSNKALPAPIETKISK 123
Db 227 PSSGK--SAVQGPPELDLCCGYSVSVLPGCAQPMWHGETFTCTAAHPELTKTPLTANITK 284

QY 124 AKGPREPQYVTLPPSRDELTKNQ-VSLTCLVKGFYPSDIAVWESNGQ--PENNYKTTTP 180
Db 285 S-GNTRPEVHLLPPSEELALNELYTLCLARGSPKDVLVWLGSGQELPREKYLTTWA 343

QY 181 PVLVD-SDG--SFPLYSKLTVDKSRWQGNVFCVSVMEALHNYTKSLSPGK 232
Db 344 SROEPSQGTTFATVTSILRVAEDWKGDFTSCMWGHEALPLAFTQKTIDRLAGK 398

RESULT 5
ID Q9QX57 PRELIMINARY; PRT; 509 AA.
AC Q9QX57;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE BIT.
DE BIT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Sano S., Ohnishi H., Kubota M.;
RT "Gene structure of mouse BIT/SHPS-1.";
RL Biochem. J. 344:667-675(1999).
DR EMBL; AB024507; BAA89289.1; JOINED.
DR EMBL; AB024500; BAA89289.1; JOINED.
DR EMBL; AB024501; BAA89289.1; JOINED.
DR EMBL; AB024502; BAA89289.1; JOINED.
DR EMBL; AB024503; BAA89289.1; JOINED.
DR EMBL; AB024504; BAA89289.1; JOINED.
DR EMBL; AB024505; BAA89289.1; JOINED.
DR EMBL; AB024506; BAA89289.1; JOINED.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; ig. 3.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 509 AA; 55997 MW; 55023F2FC1A002B5 CRC64;

```

```

Query Match 14.5%; Score 183; DB 11; Length 509;
Best Local Similarity 25.7%; Pred. No. 1e-08;
Matches 66; Conservative 42; Mismatches 107; Indels 42; Gaps 12;

QY 5 CDKTHCPCPAPPELL--GGPSVFLF-----PPK---PKDTLMISRTPEVTCVVVDVSH-- 53
Db 121 CVRFQKGPSPDTEIQGGTEVVLAKPSPVSGPADRGIDQKVNFTC----KSHGF 176

QY 54 EDEVFNVTVDGVEVHNKTKPREQYNSTYRVSVLTVLHODWLNKREYKCKVSNKAL 113
Db 177 SPRLNLTWFKDQGLHLEHTVTPNSGKNVSNISSTVRVY----LNSMDVHSGKICEVA 232

QY 114 PAPIETISKAKGPREPQVYTLPPS-----RDELTKNOVSLTCLVKGFYPSDIAVWES 168
Db 233 HTFLDR--SPRLGIANLSNFIKRVPTVKYVQQSPTSMNQVNLTCRAERFPEDLQLIWE 290

QY 169 NGPNENYKTPVPL--DSGDSFLYSLKTVDKSRWQGNVFCVSVMEALHNYTKSLSPGK 222
Db 291 NGVNSRN--DTPNLTNTDGTNTYSLFLVNSSAHREDVVFCTQVKHDDQQAIPTRNHTV 348

QY 223 -----QKSLSPG 231
Db 349 LGLAHSSDOGSMQTFFG 365

```

```

RESULT 6
ID P97797 PRELIMINARY; PRT; 513 AA.
AC P97797; P97796; O35924; O88555; O88556;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE SUBSTRATE 1 PRECURSOR
DE (P84) (SHP SUBSTRATE-1) (INHIBITORY RECEPTOR SHPS-1) (SHPS-1) (SHPS1).
GN PTPNS1 OR P84 OR SHPS-1 OR SHPS1 OR BIT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=97223399; PubMed=9070220;
RA Yamao T., Matozaki T., Amano K., Matsuda Y., Takahashi N., Ochi F.,
RA Fujioaka Y., Kasuga M.;
RT "Mouse and human SHPS-1: molecular cloning of cDNAs and chromosomal
RT localization of genes.";
RL Biochem. Biophys. Res. Commun. 231:61-67(1997).
RN [2]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 32-53; 422-424 AND 429-433.
RC STRAIN=BALB/C; TISSUE=CEREBELLUM, AND BRAIN;
RA Comu S., Weng W., Olinisky S., Ishwad P., Mi Z., Hempel J., Watkins S.,
RA Lagenaur C.F., Narayanan V.;
RT "The murine p84 neural adhesion molecule is SHPS-1, a member of the
RT phosphatase-binding protein family.";
RL J. Neurosci. 17:8702-8710(1997).
RN [3]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6; TISSUE=FETAL THYMUS;
RX MEDLINE=98380500; PubMed=9712903;
RA Veillette A., Thibaudau E., Latour S.;
RT "High expression of inhibitory receptor SHPS-1 and its association
RT with protein tyrosine phosphatase SHP-1 in macrophages.";
RL J. Biol. Chem. 273:22719-22728(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Sano S., Ohnishi H., Kubota M.;
RT "Gene structure of mouse BIT/SHPS-1.";
RL Biochem. J. 344:667-675(1999).
RN [5]
RP FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX MEDLINE=90152134; PubMed=2303162;
RA Chuang W., Lagenaur C.F.;
RT "Central nervous system antigen P84 can serve as a substrate for
RT neurite outgrowth.";
RL Dev. Biol. 137:219-232(1990).
CC -1- FUNCTION: BINDS TO THE SH2 DOMAINS OF PROTEIN-TYROSINE PHOSPHATASE
CC 2C (SHP-2) AFTER TYROSINE PHOSPHORYLATION INDUCED BY VARIOUS
CC MITOGENS AND CELL ADHESION. PROTEIN SUPPORTS ADHESION OF
CC CEREBELLAR NEURONS, NEURITE OUTGROWTH AND GLIAL CELL ATTACHMENT.
CC MAY PLAY A KEY ROLE IN INTRACELLULAR SIGNALING DURING
CC SYNAPTONEGESIS AND IN SYNAPTIC FUNCTION. MAY ACT AS A DOCKING
CC PROTEIN AND INDUCE TRANSLOCATION OF SHP-2 FROM THE CYTOSOL TO THE
CC PLASMA MEMBRANE. TYROSINE PHOSPHORYLATED PTPNS1 FROM MACROPHAGES
CC PRIMARILY ASSOCIATES WITH SHP-1.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: THREE FORMS; ISOFORMS 1 (SHOWN HERE), 2 AND
CC SMALL, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE CEREBRAL CORTEX, CEREBELLUM,
CC SPINAL CORD, SPLEEN AND MACROPHAGES AND AT MUCH LOWER LEVEL IN THE
CC KIDNEY, HEART, LIVER AND THYMUS. WITHIN THE CEREBELLUM, EXPRESSION
CC IS DETECTED IN THE SYNAPTIC GLOMERULI, GRANULE CELL BODIES, GOLGI
CC EPITHELIAL CELLS, OLFACTORY BULB, NEURONS IN THE HIPPOCAMPUS, AND
CC THE DENTATE GYRUS. IN THE EYE, EXPRESSION IS FIRST SEEN IN THE
CC INNER PLEXIFORM AND OPTIC FIBER LAYERS. LATER IN DEVELOPMENT

```

EXPRESSION IS ALSO SEEN IN THE OUTER PLEXIFORM, RETINA AND OUTER SEGMENTS OF THE PHOTORECEPTOR LAYER.

-!- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7. WHEN EXPRESSION IS RESTRICTED TO THE FLOOR PLATE REGION OF THE VENTRAL NEURAXIS. WIDESPREAD EXPRESSION IN A VARIETY OF CNS REGIONS BEGINS 2-5 DAYS AFTER BIRTH.

-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS TWO C1-LIKE AND ONE V-LIKE DOMAINS.

EMBL: D87968; BAA13521.1; -
EMBL: D87967; BAA13520.1; -
EMBL: U89694; AAB92591.1; -
EMBL: AF072543; AAC24886.1; -
EMBL: AF072544; AAC24887.1; -
EMBL: AB024507; BAA89290.1; -
EMBL: AB024500; BAA89290.1; JOINED.
EMBL: AB024501; BAA89290.1; JOINED.
EMBL: AB024502; BAA89290.1; JOINED.
EMBL: AB024503; BAA89290.1; JOINED.
EMBL: AB024504; BAA89290.1; JOINED.
EMBL: AB024505; BAA89290.1; JOINED.
EMBL: AB024506; BAA89290.1; JOINED.
MGI: 108563; Ptpnsl.
INTERPRO: IPR003006; -
PFAM: PF00047; Ig; 3.
PROSITE: PS00290; IG_MHC; UNKNOWN_1.
Signal; Transmembrane; Alternative splicing; Immunoglobulin domain; Glycoprotein; SH3-binding; Phosphorylation.

FT SIGNAL 1 31
FT CHAIN 32 513
FT
FT
FT DOMAIN 32 373 PROTEIN TYROSINE PHOSPHATASE, NON-
FT TRANSMEM 374 394 RECEPTOR TYPE SUBSTRATE 1.
FT DOMAIN 395 513 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 51 124 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 167 234 IG-LIKE C1-TYPE DOMAIN.
FT DOMAIN 270 336 IG-LIKE C1-TYPE DOMAIN.
FT DOMAIN 440 443 SH2-BINDING (POTENTIAL).
FT DOMAIN 450 456 SH2-BINDING (POTENTIAL).
FT DOMAIN 464 467 SH2-BINDING (POTENTIAL).
FT DOMAIN 481 484 SH2-BINDING (POTENTIAL).
FT DOMAIN 505 508 SH2-BINDING (POTENTIAL).
FT MOD_RES 423 423 PHOSPHORYLATION (BY CAMP) (POTENTIAL).
FT MOD_RES 440 440 PHOSPHORYLATION (BY TYR-KINASES) (POTENTIAL).
FT MOD_RES 464 464 PHOSPHORYLATION (BY TYR-KINASES) (POTENTIAL).
FT MOD_RES 481 481 PHOSPHORYLATION (BY TYR-KINASES) (POTENTIAL).
FT MOD_RES 505 505 PHOSPHORYLATION (BY TYR-KINASES) (POTENTIAL).
FT CARBOHYD 54 54
FT CARBOHYD 92 92
FT CARBOHYD 168 168
FT CARBOHYD 180 180
FT CARBOHYD 205 205
FT CARBOHYD 209 209
FT CARBOHYD 246 246
FT CARBOHYD 271 271
FT CARBOHYD 293 293
FT CARBOHYD 302 302
FT CARBOHYD 312 312
FT CARBOHYD 320 320
FT CARBOHYD 345 345
FT CARBOHYD 367 367
FT CARBOHYD 147 364
FT VARSPLIC 425 428 MISSING (IN ISOFORM SMALL).
FT CONFLICT 10 10 MISSING (IN ISOFORM 2).
FT CONFLICT 29 29 R -> L (IN REF. 2).
FT CONFLICT 67 67 V -> A (IN REF. 2 AND 3).
FT CONFLICT 74 74 K -> R (IN REF. 3).
FT CONFLICT 67 74 Q -> P (IN REF. 3).
FT CONFLICT 83 83 T -> A (IN REF. 3).
FT CONFLICT 86 87 HF -> YV (IN REF. 3).
FT CONFLICT 90 90 V -> I (IN REF. 3).

FT CONFLICT 91 91 T -> R (IN REF. 2 AND 3).
FT CONFLICT 96 96 A -> T (IN REF. 2 AND 3).
FT CONFLICT 114 114 E -> A (IN REF. 3).
FT CONFLICT 118 118 T -> I (IN REF. 3).
FT CONFLICT 126 127 KG -> RG (IN REF. 2).
FT CONFLICT 128 128 P -> S (IN REF. 2 AND 3).
FT CONFLICT 194 194 H -> P (IN REF. 3).
FT CONFLICT 224 224 H -> N (IN REF. 3).
FT CONFLICT 351 351 L -> F (IN REF. 3).
FT CONFLICT 365 365 G -> D (IN REF. 3).
FT CONFLICT 490 490 S -> N (IN REF. 2).
SQ SEQUENCE 513 AA; 56424 MW; 2AFDD5FA6C1EC7C CRC64;

Query Match 14.5%; Score 183; DB 11; Length 513;
Best Local Similarity 25.7%; Pred. No. le-08;
Matches 66; Conservative 42; Mismatches 107; Indels 42; Gaps 12;

QY 5 CDKTHTCPCPAPPELL--GGPSVFLF-----PPK---PKDTLMISRTPEVTCVVVDVSH-- 53
DB 121 CVKQKGPSEPDTEIQSGGTEVIVLAKPSPPEVSGPADRGIPDKVNFCTC-----KSHGF 176
QY 54 EDPEVKFMYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKAL 113
DB 177 SPRNITLKFQKDGQELHLETTVPNSGKNVSYNISTVRV-----LNSMDVHSKVCEVA 232
QY 114 PAPIETKTSKAGQPREPQVVTLPSS-----RDELTKNQSVLTCLVKGFYPSDIATVWES 168
DB 233 HITLDR--SPLRIGIANLSNFRVSVTKVTKQOSTPSMNVNLTCTRAERFYEDQLQLWLE 290
QY 169 NGQPENNYKTTTPPVL--DSGGSFFLYSKLTVDKSRWQGNVFCGSMHME---ALHNHYT- 222
DB 291 NGNVSRL--DTPKNLTKNVTDGTYNTSLFLYNSSAHRDVVFTCVQKHQQPATRNHTV 348
QY 223 -----OKSLSLSPG 231
DB 349 LGLAHSDDQSGSMQTFFG 365

RESULT 7
Q90524 PRELIMINARY; PRT; 268 AA.
AC Q90524;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR (FRAGMENT).
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]
RP SEQUENCE OF 1-240 FROM N.A.
RC TISSUE-SPLEEN;
RX MEDLINE=95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes M., Hughes E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks.";
RL Nature 374:168-173(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RA Greenberg A.S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U18680; AAB48352.1; -
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; Ig; 2.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 268 268
SQ SEQUENCE 268 AA; 29579 MW; 0CDE524DD6E9FDC4 CRC64;

Query Match 14.3%; Score 181; DB 13; Length 268;
Best Local Similarity 25.1%; Pred. No. 7e-09; Indels 60; Gaps 9;
Matches 62; Conservative 35; Mismatches 90; Indels 60; Gaps 9;

QY 23 PSVFLPPKPKDTLMISRTPE-----VTCVVVDVSHEDPEVKFNWYVDG----- 66
DB 14 PYVFLPPKPKDTLMISRTPE-----TARVDQTPQETTKETGSGSLNCALPSTYWNKKSGSTNEE 65
QY 67 -----VEVINAKTKPREEQNSTYRVVSVLTVLHODWLNKGEYKCKV--SNKALPAP 116
DB 66 TTSKGRVYVETVNSGKSFSLRIND-----LTV-----EDSGTYRCVKYRKNNAAYDCG 113
QY 117 IEKT-----ISKAKGQPREPQVYTL--PPSRDELTKNOVSLTCLVKGFPSPDIA 163
DB 114 LEELDWIYVGGGTGVNPGIPLSPPIVSLLSHSAEEORANGFVOLVCLISGYYPENIA 173
QY 164 VEWESNGQP-ENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYT 222
DB 174 VSWQNTKTIISGFATTSVTKTSSNDFSCASLLKVLQEWMSRGSVYSCQVSHSATSSNQ 233
QY 223 QKSLSLS 229
DB 234 KEIRSTS 240

RESULT 8
Q46632 PRELIMINARY; PRT; 506 AA.
AC O46632;
DT 01-JUN-1998 (TREMBlrel.. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE MID-1 ANTIGEN PRECURSOR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98143722; PubMed=9485180;
RA Brooke G.P., Parsons K.R., Howard C.J.;
RT "Cloning of two members of the SIRP alpha family of protein tyrosine
phosphatase binding proteins in cattle that are expressed on monocytes
and a subpopulation of dendritic cells and which mediate binding to
CD4 T cells";
RL Eur. J. Immunol. 28:1-11(1998).
DR EMBL; Y11045; CAA71942.1;
DR INTERPRO; IPR003006;
DR PFAM; PF00047; ig; 3.
KW Signal.
FT CHAIN 1 26 POTENTIAL.
FT CHAIN 27 506 MYD-1 ANTIGEN.
SQ SEQUENCE 506 AA; 55117 MW; ACDDEA9C9708A82 CRC64;

Query Match 14.2%; Score 179; DB 6; Length 506;
Best Local Similarity 27.0%; Pred. No. 2.3e-08;
Matches 62; Conservative 36; Mismatches 108; Indels 24; Gaps 6;

QY 1 EPKSCDKTH-TCPCCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSH--EDPE 57
DB 133 EFKSGCGTHLVNAKSPVPPVSGPTV-----RATPEQVNFCTSHGSPRN 179
QY 58 VRFNWKYVDGVEVHNKTKPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPA-- 115
DB 180 ISLWFKNGKNLSASQTSVDPEDDNDYSINSTTKVLLATGDVHSQVCEVAHVTLQGGP 239
QY 116 PIBKTKSKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPENN 175
DB 240 PLRGANLSETIRVPP---TLEITRSPSAGNQNVNVCQNKFPYPRHLQLTLWLENGNMSRT 296

QY 176 YKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHE---ALHNHYT 222
DB 297 EAASVLVENKDGTFNQTSWLLVNSSAHRVAVLTQCVHEDRQPAVSKNHT 346

RESULT 9
O08907 PRELIMINARY; PRT; 509 AA.
ID O08907;
AC O08907;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE BRAIN IMMUNOLOGICAL-LIKE PROTEIN (BIT).
GN BIT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=BRAIN;
RX MEDLINE=97230468; PubMed=9073522;
RA Ohnishi H., Kubota M., Sano S.;
RT "Bit (Bit) maps to mouse chromosome 2.";
RL Genomics 40:504-506(1997).
DR EMBL; D85785; BAA20376.1;
DR MGD; MGI:107947; Bit.
DR INTERPRO; IPR003006;
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN.1.
SQ SEQUENCE 509 AA; 56033 MW; 1E377ACEC31BAL3E CRC64;

Query Match 14.0%; Score 177; DB 11; Length 509;
Best Local Similarity 27.7%; Pred. No. 3.6e-08;
Matches 66; Conservative 38; Mismatches 96; Indels 38; Gaps 12;

QY 21 GGSVFLF---PPK---PKDTLMISRTPEVTCVVVDVSH--EDPEVKFNWYVDGVEVHN 71
DB 139 GGTEVYVLAKPSPPEVSGPADRGIPDKVNFTC-----KSHGFSFRNITLKWFGDQGLHP 194
QY 72 AKTKPREEQNSTYRVVS-VLTVLHODWLNKGEYKCKVSNKALPAPIETKISKAKGQPR- 129
DB 195 LETVNPSPGNVSNYSSITVVRVVLNSMDVNSKVI-CEVAHITLDRSLRGIANLSNIRV 253
QY 130 --EPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPENNYKTPPVL--DS 185
DB 254 SPTVKVYTPQPPPT---SMNQVNLTCRAERFYPEDLQILWLENGNVSRN--DTPKNLTNT 307
QY 186 DGSFFLYSKLTVDKSRWQGNVFCSCVMHE---ALHNHYT-----OKSLSLSPG 231
DB 308 DGYNTSLFLVNSSAHRVDVFTQVKKHQDQPAITRNHTVLGLAHSSDQSGMTFFG 365

RESULT 10
O46631 PRELIMINARY; PRT; 506 AA.
ID O46631;
AC O46631;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE MYD-1 ANTIGEN PRECURSOR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PERIPHERAL BLOOD;
RX MEDLINE=98143722; PubMed=9485180;
RA Brooke G.P., Parsons K.R., Howard C.J.;
RT "Cloning of two members of the SIRP alpha family of protein tyrosine

DT 01-OCT-2000 (TREMELREL. 15, Last sequence update)
DE 01-OCT-2000 (TREMELREL. 15, Last annotation update)
DE MHC CLASS II ALPHA SUBUNIT.
GS MHC-AUHA-DBA.

OS Aulonocara hansbaenschi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthopterygii; Acanthopterygii; Perciformes; Labroidel;
OC Cichlidae; Aulonocara.
OX NCBL_TaxID=27781;

RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Suelmann H., Klein J.;
RT "Major histocompatibility complex class II A genes in cichlid fishes:
RT identification, expression, linkage relationships, and haplotype
RT variation";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF212849; AAF5676.1; -.
KW MHC.

SQ SEQUENCE 238 AA; 26005 MW; 3D18BF05BF0A9C53 CRC64;

Query Match 13.4%; Score 169; DB 7; Length 238;
Best Local Similarity 29.9%; Pred. No. 7.6e-08;
Matches 63; Conservative 27; Mismatches 83; Indels 38; Gaps 10;

QY 44 VTCVVVDVSHEDPEV-----KENWYVDGVEVHNATKPREEQYNSTYRVSVLTVLH 95
DB 11 VLVCSADVHLHDIQVIGSDSDGEMYGIDGEEKAYADENKKEEIIYQPPFVNPFP--Y 68
QY 96 QDWLNGKEYCKVCKVSNKALPAPIETKISKAKGQPREP-----QVYTLPPSRDEL---TK 145
DB 69 QE-----GAYENAVANQICRENKRVKGMKDYPLEQVAPSGVMYIT-----RDEVEFGK 120
QY 146 NOVSLTCLVKGYPSPDI AVEVESNGQPPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOG 205
DB 121 N---TLCHVTGTPAPVNVSWTKNGQKVTGSSINTYPNKGDTFTQISRLQFTP---QLG 175
QY 206 NVFSGSVMEALHNHYTQ-----KSLSLSPG 231
DB 176 DIYSCAVQHATQATKFEYVDSSAQSDPG 206

RESULT 14
QY0536 PRELIMINARY; PRT; 257 AA.
AC QY0536;

DT 01-NOV-1996 (TREMELREL. 01, Created)
DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
DT 01-OCT-2000 (TREMELREL. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR (FRAGMENT).
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBL_TaxID=7801;

RN [1]
RP SEQUENCE OF 1-242 FROM N.A.

RC TISSUE-SPLEEN;
RX MEDLINE=95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;

RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks.";
RL Nature 374:168-173(1995).
RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RA Greenberg A.S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U18693; AAB48364.1; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; 19; 2.

DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 257
SQ SEQUENCE 257 AA; 28083 MW; BAD019A99D0BC491 CRC64;

Query Match 13.2%; Score 166.5; DB 13; Length 257;
Best Local Similarity 29.4%; Pred. No. 1.4e-07;
Matches 40; Conservative 21; Mismatches 64; Indels 11; Gaps 3;

QY 97 DWLNGKEYCKVCKVSNKALPAPIETKISKAKGQPREPQVYTL--PPSRDELTKNOVSLTCLV 154
DB 115 DWAVGGAYAAVGDTA-----VTVNPFGIPSPPIVLLHSATEEQRRANREVQLVCLI 166
QY 155 KGFYPSDIAVEVESNGQPPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOGNVEFSCSYM 213
DB 167 SGYPENIAVSWQKNTKTITSGFATTPSVKTSNDFSCASLLKVPLOEWSRGVSQCQVS 226
QY 214 HEALHNHYTQKLSLS 229
DB 227 HSATSSNQKKEIRSTS 242

RESULT 15
QY0545 PRELIMINARY; PRT; 237 AA.
AC QY0545;

DT 01-NOV-1996 (TREMELREL. 01, Created)
DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
DT 01-OCT-2000 (TREMELREL. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR (FRAGMENT).
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBL_TaxID=7801;

RN [1]
RP SEQUENCE OF 1-219 FROM N.A.

RC TISSUE-SPLEEN;
RX MEDLINE=95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;

RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks.";
RL Nature 374:168-173(1995).
RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-SPLEEN;
RA Greenberg A.S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U18702; AAB48196.1; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; 19; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 237
FT NON_TER 237
SQ SEQUENCE 237 AA; 25831 MW; 9C002BD0B55C8656 CRC64;

Query Match 13.1%; Score 165; DB 13; Length 237;
Best Local Similarity 25.8%; Pred. No. 1.8e-07;
Matches 56; Conservative 32; Mismatches 83; Indels 46; Gaps 8;

QY 44 VTCVVVDV-----SHEDPEVKNWYVDGVEVHNATKPREEQYNSTYRV 87
DB 17 INCVLRDATSALGNTCWSRKKSLSTREETIAGRGYVETV---NSGSKSFSLIND--- 69
QY 88 VSVLTVLHODWLNKEYCKVCKVSN-KALPAPIETKISKAKGQPREPQVYTL 135
DB 70 ---LTV-----EDGTYRCGVGVYRCPCFCLSHLSYAGCGDGTVVTVNFPPIVPS 121
QY 136 L--PPSRDELTKNOVSLTCLVKGYPSPDI AVEVESNGQPPENNYKTPPVLDSDGSFFLY 192
DB 122 LLHSATEEQRRANREVQLVCLISGYYPENIAVSWQKNTKTITSGFATTPSVKTSNDFSCA 181

Qy 193 SKLTVDKSRWQGNVFCVNMHEALHNYTQKSLSL 229
| | | | | : | : | | | : : | |
Db 182 SLLKVPLEWGRGVSQVSHSATSSNQKEIRSTS 218

Search completed: March 1, 2001, 09:20:02
Job time: 400 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:13:22 ; Search time 40.97 Seconds
(without alignments)
182.871 Million cell updates/sec

Title: US-09-389-782A-1

Perfect score: 1263

Sequence: 1 EPKCDKTHTPCPAPPELL.....MHEALHNHYTKLSLSPGK 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1263	100.0	330	1 GCL_HUMAN	P01857 homo sapien
2	1186	91.5	290	1 GC3_HUMAN	P01860 homo sapien
3	1145	90.7	326	1 GC2_HUMAN	P01859 homo sapien
4	1135	89.9	327	1 GC4_HUMAN	P01861 homo sapien
5	921	72.9	323	1 GC_RABIT	P01870 oryctolagus
6	896	70.9	329	1 GC2_CAYPO	P01862 cavia porce
7	845.5	66.9	329	1 GC3_MOUSE	P22436 mus musculus
8	842	66.7	333	1 GCB_RAT	P20761 rattus norv
9	834.5	66.1	398	1 GC3M_MOUSE	P03987 mus musculus
10	822.5	65.1	324	1 GCL_MOUSE	P01868 mus musculus
11	822.5	65.1	326	1 GCL_RAT	P20759 rattus norv
12	817.5	64.7	393	1 GC1M_MOUSE	P01869 mus musculus
13	809.5	64.1	329	1 GCC_RAT	P20762 rattus norv
14	809	64.1	330	1 GCRA_MOUSE	P01863 mus musculus
15	805	63.7	335	1 GCAB_MOUSE	P01864 mus musculus
16	804	63.7	399	1 GCAM_MOUSE	P01865 mus musculus
17	785.5	62.2	322	1 GCA_RAT	P20760 rattus norv
18	779.5	61.7	336	1 GCB_MOUSE	P01866 mus musculus
19	774.5	61.3	405	1 GCEN_MOUSE	P01867 mus musculus
20	358	28.3	454	1 MUC_HUMAN	P01871 homo sapien
21	357	28.3	455	1 MUC_MOUSE	P01872 mus musculus
22	356	28.2	391	1 MUCB_HUMAN	P04220 homo sapien
23	354	28.0	421	1 EPC_MOUSE	P06336 mus musculus
24	353	27.9	429	1 EPC_RAT	P01855 rattus norv
25	352	27.9	428	1 EPC_HUMAN	P01854 homo sapien
26	347	27.5	476	1 MUCN_MOUSE	P01873 mus musculus
27	345	27.3	458	1 MUC_RABIT	P03988 oryctolagus
28	335	26.5	479	1 MUCM_RABIT	P04221 oryctolagus
29	332	26.3	450	1 MUC_CANFA	P01874 canis famil
30	332	26.3	454	1 MUC_MESAU	P06337 mesocricetu
31	329	26.0	457	1 MUC_SUNNO	P20768 suncus muri
32	297	23.5	438	1 HVC2_HETFR	P23085 heterodontu
33	295.5	23.4	299	1 ALC_RABIT	P01879 oryctolagus

ALIGNMENTS

RESULT 1
GCL_HUMAN
ID GCL_HUMAN STANDARD; PRT; 330 AA.
AC P01857;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE IG GAMMA-1 CHAIN C REGION.

GN IGHG1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=82274238; PubMed=6287432;

RA Ellison J.W., Berson B.J., Hood L.E.;

RT "The nucleotide sequence of a human immunoglobulin C gamma gene."

RL Nucleic Acids Res. 10:4071-4079(1982).

RN [2]

RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).

RX MEDLINE=71064024; PubMed=5489771;

RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,

Waxdal M.J., Edelman G.M.;

RT acid sequence of heavy-chain of a human gamma G-immunoglobulin. VII. Amino

RL Biochemistry 9:3161-3170(1970).

RN [3]

RP SEQUENCE OF 136-329 (EU).

RX MEDLINE=71064025; PubMed=5530842;

RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,

Edelman G.M.;

RT acid sequence of heavy-chain of a human gamma G-immunoglobulin. 8. Amino

RL Biochemistry 9:3171-3181(1970).

RN [4]

RP SEQUENCE (MYELOMA PROTEIN NIE).

RX MEDLINE=77070269; PubMed=826475;

RA Ponstingl H., Hilschmann N.;

RT "The rule of antibody structure. The primary structure of a

monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The

chymotryptic peptides of the H-chain, alignment of the tryptic

peptides and discussion of the complete structure."

RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).

RN [5]

RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.

RX MEDLINE=83289131; PubMed=6884994;

RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;

RT "Three-dimensional structure determination of antibodies. Primary

structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."

RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).

RN [6]

RP DISULFIDE BONDS.

RX MEDLINE=71064027; PubMed=4923144;

RA Gall W.E., Edelman G.M.;

RT "The covalent structure of a human gamma G-immunoglobulin. X.

RL Intrachain disulfide bonds."

RT

Biochemistry 9:3188-3196(1970).
 [7]
 RN DISULFIDE BONDS.
 RX MEDLINE=77070267; PubMed=1002129;
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT cyanogen bromide cleavage products, and the disulfide bridges.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 [8]
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RP MEDLINE=81208100; PubMed=7236608;
 RA Deisenhofer J.;
 RT "Crystallographic refinement and atomic models of a human Fc fragment
 RT and its complex with fragment B of protein A from Staphylococcus
 RT aureus at 2.9- and 2.8-A resolution.";
 RL Biochemistry 20:2361-2370(1981).
 CC -!- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE
 CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)
 CC MARKER & THE GIM (NON-1) MARKERS.
 CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
 CC 35,116,198,269 & 272.
 CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
 CC 268-272.
 CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
 CC RESIDUES 198,267&272.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; J00228; AAC82527.1; ALT_INIT.
 DR PIR; A02146; GHU.
 DR PDB; 1FC1; 15-JUL-92.
 DR PDB; 1FC2; 15-JUL-92.
 DR MIM; 147100; -.
 DR INTERPRO; IPR000495; -.
 DR INTERPRO; IPR003006; -.
 DR PFAM; PF00047; 1g; 3.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 3D-structure.
 FT NON_TER 1 1
 FT DOMAIN 1 98
 FT HINGE 1 98
 FT CH1.
 FT HINGE.
 FT CH2.
 FT CH3.
 FT DISULFID 27 83
 FT DISULFID 103 103
 FT DISULFID 109 109
 FT DISULFID 112 112
 FT DISULFID 144 204
 FT DISULFID 250 308
 FT CARBOHYD 180 180
 FT VARIANT 97 97
 FT VARIANT 239 239
 FT VARIANT 241 241
 FT MOD_RES 330 330
 FT STRAND 123 126
 FT HELIX 130 134
 FT TURN 136 137
 FT STRAND 141 148
 FT STRAND 158 162
 FT TURN 163 164

FT STRAND 165 166
 FT STRAND 175 178
 FT STRAND 183 190
 FT HELIX 193 197
 FT TURN 198 199
 FT STRAND 202 206
 FT STRAND 215 219
 FT STRAND 227 227
 FT STRAND 230 234
 FT HELIX 238 240
 FT TURN 241 242
 FT STRAND 245 256
 FT STRAND 260 266
 FT TURN 267 268
 FT STRAND 269 270
 FT STRAND 274 276
 FT STRAND 280 281
 FT TURN 283 284
 FT STRAND 287 296
 FT HELIX 297 301
 FT TURN 302 303
 FT STRAND 306 312
 FT TURN 313 314
 FT TURN 316 317
 FT STRAND 320 324
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;
 Query Match 100.0%; Score 1263; DB 1; Length 330;
 Best Local Similarity 100.0%; Pred. No. 3.2e-95;
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPKSCDTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
 DB 99 EPKSCDTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKF 158
 QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120
 DB 159 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 218
 QY 121 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 180
 DB 219 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 278
 QY 181 PVLSDSGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTQKLSLSLSPCK 232
 DB 279 PVLSDSGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTQKLSLSLSPCK 330
 RESULT 2
 GC3_HUMAN
 ID GC3_HUMAN STANDARD; PRT; 290 AA.
 AC P01860;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).
 GN IGHG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE (DISEASE PROTEIN WIS).
 RX MEDLINE=81021548; PubMed=6774747;
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
 RT gamma 3 heavy-chain disease protein Wis.";
 RL Biochemistry 19:4304-4308(1980).
 RN [2]
 RP REVISIONS TO 12-97 OF PROTEIN WIS.
 RX MEDLINE=77118561; PubMed=402363;
 RA Michaelisen T.E., Frangione B., Franklin E.C.;
 RT "Primary structure of the 'hinge' region of human IgG3. Probable

Best Local Similarity 93.7%; Pred. No. 7.1e-85;
Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 11 CPGCPAPPELLGSPVFLFPKPKDTLMISRTPEVTVVVDVSHEDPEVFNWYVDGVEVH 70
Db 106 CPSCPAPELGGSPVFLFPKPKDTLMISRTPEVTVVVDVSDQEDPEVFNWYVDGVEVH 165

QY 71 NAKTPREEQNSTYRVSVLVTHODWLNKGYKCKVSNKALPAPIEKTISKAKQPRE 130
Db 166 NAKTPREEQNSTYRVSVLVTHODWLNKGYKCKVSNKALPAPIEKTISKAKQPRE 225

QY 131 PQVYTLPPSRDELTKNQVSLTCLVKGYFSPDSIAVESNGQPENNYKTPPVLDSDGSEFF 190
Db 226 PQVYTLPPSRDELTKNQVSLTCLVKGYFSPDSIAVESNGQPENNYKTPPVLDSDGSEFF 285

QY 191 LYSKLTVDKSRWQGNVSCSVNHEALHNHYTQKSLSLSPGK 232
Db 286 LYSRLTVDKSRWQGNVSCSVNHEALHNHYTQKSLSLSPGK 327

RESULT 5
GC_RABBIT STANDARD; PRT; 323 AA.
ID GC_RABBIT
AC P01870;
DF 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA CHAIN C REGION.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RT F-1 haplotype."
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RT immunoglobulin G of different allotype."
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RT heavy chain and identification of two genomic C gamma genes."
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RT immunoglobulin G."
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -!- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
entties requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL; M16426; AAA31289.1; -.
CC PIR; A02161; GHRB.
CC INTERPRO; IPR000495; -.
CC INTERPRO; IPR003006; -.
CC PFAM; PF00047; ig; 3.
CC PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> E (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).
FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 72.9%; Score 921; DB 1: Length 323;
Best Local Similarity 67.3%; Pred. No. 1.4e-67;
Matches 167; Conservative 31; Mismatches 34; Indels 16; Gaps 2;

QY 1 EPKSCDKTH-----TC--PPCAPPELLGGPSVFLFPKPKDTLMISRTPEV 44
Db 76 QPVTCTNVAHPATNTKVDKTVAPSTCSKTPCPPELLGGPSVFLFPKPKDTLMISRTPEV 135

QY 45 TCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEY 104
Db 136 TCVVVDVSDQEDPEVFQVTVYINNEQVRTAPRPLRQDFNFTSTVRLPITHQDWLNGKEY 195

QY 105 KCKVSNKALPAPIEKTISKAKGQPREPPQVYTLPPSRDELTKNQVSLTCLVKGYFSPDSIAV 164
Db 196 KCKVSNKALPAPIEKTISKAKGQPREPPQVYTLPPSRDELTKNQVSLTCLVKGYFSPDSISV 255

QY 165 EWESNGQPNENNYKTPPVLDSDGSEFFLYSKLTVDKSRWQGNVSCSVNHEALHNHYTQK 224
Db 256 EWEKNGKAEDNYKTPPVLDSDGSEFFLYSKLTVDKSRWQGNVSCSVNHEALHNHYTQK 315

QY 225 SLSLSPGK 232
Db 316 SISRSFGK 323

RESULT 6
GC2_CAVPO STANDARD; PRT; 329 AA.
ID GC2_CAVPO
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2 CHAIN C REGION.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.

```
RX MEDLINE=71058471; PubMed=5538606;
RA Birshtein B.K., Hussain Q.Z., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the
RT half-cystine joining heavy and light chains.";
RL Biochemistry 10:18-25(1971).
RN [3]
RN SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
RT and hinge region cyanogen bromide fragments.";
RL Biochemistry 10:9-17(1971).
RN [4]
RN SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E., Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4796-4803(1974).
RN [5]
RN SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;
RA Trischmann T.M., Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4804-4811(1974).
RN [6]
RN DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B., Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
RL Biochemistry 10:26-31(1971).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR: A02151; G2GP.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 70.9%; Score 896; DB 1; Length 329;
Best Local Similarity 70.4%; Pred. No. 1.5e-65;
Matches 164; Conservative 25; Mismatches 38; Indels 6; Gaps 2;

QY 1 EPKSCDTHTCPPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60
Db 101 ZPBPC-----TCPKPPENLGGPSVFIFFPKPKDTLMISLTPRVTCTVVDVSDPEVQF 156

QY 61 NWYDGVHNAKTPREQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKG 120
Db 157 TWFDVNDKPVGNAETKPRVEQNTTFVESVLPVQHDWLRCGEKCKVYNKALPAPIEKT 216

QY 121 ISKAGQPEQVYTLPPSRDELTKNVSCLVKGYFSPDAVWEWSNGQPF--ENNYKT 178
Db 217 ISKTGKGRMPDQVYTLPPSRDELSKSVTCLINFPDPAIHVEWASNRVPSVEKEYKN 276

QY 179 TPPVLDSGDSFFLYSKLTVDKSRWQGNVFSCVYHMEALHNHYTKSLSPG 231
Db 277 TPIEDADGSGFYLYSKLTVDKSAWDQGTQVYTCVYHMEALHNHYTKSLSPG 329
```

RESULT 7

```
GC3_MOUSE
ID GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, SECRETED FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; -. NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin C region; Glycoprotein; Transmembrane;
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;
```

Query Match 66.9%; Score 845.5; DB 1; Length 329;
Best Local Similarity 67.0%; Pred. No. 1.9e-61;
Matches 150; Conservative 34; Mismatches 37; Indels 3; Gaps 1;

```
QY 12 PP---CPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGV 68
Db 106 PGSSCPPGNILGGPSVFIFFPKPKDALMISLTPKTCVVDVSEDDPDVHVSFVDNKE 165

QY 69 VHNKATPREQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKG 128
Db 166 VHTAWTPREAQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKG 25

QY 129 REPQVYTLPPSRDELTKNVSCLVKGYFSPDAVWEWSNGQPENNYKTTTPVLDSDGS 88
Db 226 QTPQVYTIPTPREQMSKKVSLTCLVTNFTFSEATSVEMRGELEQDYKNTPTILDSGDT 85

QY 189 FFYLSKLTVDKSRWQGNVFSCVYHMEALHNHYTKSLSPG 232
Db 286 YFLYSKLTVDKSRWQGNVFSCVYHMEALHNHYTKSLSPG 329
```

RESULT 8

```
GC3_MOUSE
ID GC3_MOUSE STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
```

DE IG GAMMA-2B CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family";
RL Gene 74:473-482(1988).
DR PIR: P80018; PS0018.
DR INTERPRO: IPR000495;
DR PFAM: PF00047; ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 66.7%; Score 842; DB 1; Length 333;
Best Local Similarity 65.4%; Pred. No. 3.6e-61;
Matches 149; Conservative 33; Mismatches 46; Indels 0; Gaps 0;
QY 5 CDKTHPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 64
DB 106 CPTCTCHCKPPELLGGPSVFLFPPKPKDTLLISQNAKVTCTVVVDVSEEDPVQSFWEV 165
QY 65 DGEVHNHAKTPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISK 124
DB 166 NNVVEHTAQTPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISK 225
QY 125 KGQPREQYVTLPPSRDELTKNQVSLTCLVKGYFSPDAVWESNGQPNNTYTPPVLD 184
DB 226 KGLVRKQVYVMPPTPEQLTEQTSVLTCLTSGFLPDIGVEWTSNGHIEKNYKNTPEVMD 285
QY 185 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPGK 232
DB 286 SDGSFFMYSKLNVRSRWDSPAPFCVSVVHEGLHNHVKESIRPPGK 333

RESULT 9
GC3M_MOUSE STANDARD; PRT; 398 AA.
AC F03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene

RT segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00451; AAB59655.1;
DR EMBL: V01526; CAA24767.1; ALT_SEQ.
DR PIR: A02155; G3MSM.
DR INTERPRO: IPR000495;
DR INTERPRO: IPR003006;
DR PFAM: PF00047; ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;
Query Match 66.1%; Score 834.5; DB 1; Length 398;
Best Local Similarity 66.7%; Pred. No. 1.8e-60;
Matches 148; Conservative 34; Mismatches 37; Indels 3; Gaps 1;
QY 12 PP---CPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 68
DB 106 PGSSCPGPNILGGPSVFLFPPKPKDMLSLTPKVCVVVDVSEDDPDVHVSFVNDKE 165
QY 69 VHNAKTPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAGQP 128
DB 166 VHTAWTPQPREAQNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAGRA 225
QY 129 REPQVTLPPSRDELTKNQVSLTCLVKGYFSPDAVWESNGQPNNTYTPPVLDSDGS 188
DB 226 QTPQVYTIPTPREQMSKKVSLTCLTNTFTSEAISEWERNGELEQDYKNTPTLLSDGT 285
QY 189 FFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSP 230
DB 286 YFLYSKLTVDTSWLOGEIFTCSVWHEALHNHHTQKNLSRSP 327
RESULT 10
GC1_MOUSE STANDARD; PRT; 324 AA.
AC F01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene";
RL Cell 18:559-568(1979).

[2]
RN SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE=80202559; PubMed=6769752;
RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
MAO Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
[3]
RN SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=80012837; PubMed=113776;
RA Rogers J., Clarke P., Salsler W.;
RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
heavy chain.";
RL Nucleic Acids Res. 6:3305-3321(1979).
[4]
RN SEQUENCE (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=78242288; PubMed=98524;
RA Adetudbo K.;
RT "Evolution of immunoglobulin subclasses. Primary structure of a
murine myeloma gamma chain.";
RL J. Biol. Chem. 253:6068-6075(1978).
[5]
RN DISULFIDE BONDS (MOPC 21).
RX MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; V00793; CAA24172.1; -;
DR EMBL; V00793; CAA24173.1; -;
DR EMBL; V00793; CAA24174.1; -;
DR EMBL; V00793; CAA24175.1; -;
DR EMBL; V00795; CAA24176.1; -;
DR PIR; A02159; GIMS.
DR MGD; MGI:96446; Igh-4.
DR INTERPRO; IPR000495; -;
DR INTERPRO; IPR003006; -;
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217
FT DOMAIN 218 324 CH2.
FT DISULFID 27 82 CH3.
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT MOD_RES 324 324
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 65.1%; Score 822.5; DB 1; Length 324;
Best Local Similarity 60.9%; Pred. No. 1.3e-59;
Matches 142; Conservative 45; Mismatches 37; Indels 9; Gaps 3;

QY 2 PKCDKTHTCPP--CPAPELGGSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVK 59
Db 99 PRDCG---CKPCICTVPEV---SSVFIFPKPKDVLITITLPKVTCTVVVDISKDDPEVQ 153
QY 60 FNWYVDGVEVHNATKPREQYNSYRVSVLTCLVHODWLNKGEYCKVSKNKAIPAPIEK 119
Db 152 FSWFVDDVEVHTAQTQPREQNSFTRSVSELPIMHODWLNKGEYCKVSKNKAIPAPIEK 213
QY 120 TISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIKAVESNGQSPENNYKT 179
Db 212 TISTKGRKAPQVYTLPPPKAEQAKDKVSLTCLMTDEFFPEDIIVEMQNGQPAENYKNT 271
QY 180 PVLDSGSEFELYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSLSPGK 232
Db 272 QPIMNTNGSYFYVSKLVNPKVQKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 324
RESULT 11
GCL_RAT
ID GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0017; PS0017.
DR INTERPRO; IPR000495; -;
DR INTERPRO; IPR003006; -;
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
FT CARBOHYD 176 176 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 65.1%; Score 822.5; DB 1; Length 326;
Best Local Similarity 61.5%; Pred. No. 1.3e-59;
Matches 144; Conservative 43; Mismatches 38; Indels 9; Gaps 3;

QY 2 PKCDKTHTCPPAPPELLGG---PSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEV 58
Db 99 PRNCG--GDCKPC-----ICTSEVSSVFIFPKPKDVLITITLPKVTCTVVVDISQDPEV 152
QY 59 KFNWYVDGVEVHNATKPREQYNSYRVSVLTCLVHODWLNKGEYCKVSKNKAIPAPIEK 118
Db 153 HFSWFVDDVEVHTAQTQPREQNSFTRSVSELPILHODWLNKGEYCKVSKNKAIPAPIEK 212
QY 119 KTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIKAVESNGQSPENNYKT 178
Db 213 KTISKPEGRQVPHYVMTSPKTEMTQNEVSICTMKGFPYDPDIYVEMQNGQPAENYKNT 272
QY 179 TTPVLDSDGSEFELYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSLSPGK 232

Db 273 TPPTMDTDSFLYSLNVAKKWQGNFTCSVLHGLHNHHTKSLSHSPGK 326
 RESULT 12
 GC1M_MOUSE STANDARD; PRT; 393 AA.
 AC F01869;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 RT gamma 1 chain gene.";
 RL Cell 18:559-568(1979).
 RN [2]
 RP SEQUENCE OF 323-393 FROM N.A.
 RX MEDLINE=82197626; PubMed=6804950;
 RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
 RT "mRNA for surface immunoglobulin gamma chains encodes a highly
 RT conserved transmembrane sequence and a 28-residue intracellular
 RT domain".
 RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
 RN [3]
 RP SEQUENCE OF 323-366 FROM N.A.
 RX MEDLINE=82115295; PubMed=6799207;
 RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
 RA Eisenberg D., Wall R.;
 RT "Gene segments encoding transmembrane carboxyl termini of
 RT immunoglobulin gamma chains.";
 RL Cell 26:19-27(1981).
 RN [4]
 RP SEQUENCE OF 1-44 FROM N.A.
 RX MEDLINE=8222190; PubMed=6283537;
 RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
 RT "Nucleotide sequences of gene segments encoding membrane domains of
 RT immunoglobulin gamma chains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
 CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
 CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
 CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
 CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
 CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
 CC SEGMENT OF MU CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V00793; CAA24172.1; -;
 DR EMBL; V00793; CAA24173.1; -;
 DR EMBL; V00793; CAA24174.1; -;
 DR PIR; B02159; GIMSM.
 DR MGD; MGI:96446; IGH-4.
 DR INTERPRO; IPR000495; -;
 DR INTERPRO; IPR003006; -;
 DR PFAM; PF00047; Ig; 3.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW Alternative splicing; Transmembrane.
 FT NON_TER 1 1

FT DOMAIN 1 97 CHL.
 FT DOMAIN 98 110 HINGE.
 FT DOMAIN 111 217 CH2.
 FT DOMAIN 218 324 CH3.
 FT DISULFID 27 82
 FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 138 198
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
 FT DISULFID 244 302
 FT TRANSMEM 340 357 POTENTIAL.
 FT DOMAIN 358 393 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7ALCE27 CRC64;
 Query Match 64.7%; Score 817.5; DB 1; Length 398;
 Best Local Similarity 60.8%; Pred. No. 4.2e-59;
 Matches 141; Conservative 45; Mismatches 37; Indels 9; Gaps 3;
 QY 2 PKSCDKTHTCPP--CPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVYDVSHEDPEVK 59
 Db 99 PRDCG----CKPCICVPEV---SSVFIFPKPKADVLTITLTPKVTGVVDISKDDEVQ 151
 QY 60 FNVYDGVVEVHNKTRPREQYNSTRVSVLTVLHODWLNKREYKCKVSNKALPAPIEK,119
 Db 152 FSNFVDDVEHTAQTQPREQFNSTFRSVSELPIMHQDWLNKREKFCRVNSAAPPADIEK 211
 QY 120 TISKAKQPREQVYTLPPSRDELTKNOVSLTCLVKGFPYSDTAVENESNGQPNNYKTT 179
 Db 212 TISKTGRKPAQYITIPPKPEQMAKDVKSLTCTMTITFPEDITVEWQWQPAENYKNT 271
 QY 180 PPVLDSDGSFFLYSLNVAKKWQGNFTCSVLHGLHNHHTKSLSHSPG 323
 Db 272 QPIMNTNGSYFVYSKLVNQSKEAGNTFTCSVLHGLHNHHTKSLSHSPG 323
 RESULT 13
 ID GCC_RAT STANDARD; PRT; 329 AA.
 AC P20762;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-2C CHAIN C REGION.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88166903; PubMed=3127222;
 RA Brueggemann M., Delmastro-Galfré P., Waldmann H., Calabi F.;
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
 RT region cDNA: extensive homology to mouse gamma 3.";
 RL Eur. J. Immunol. 18:317-319(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X07189; CAA30169.1; -;
 DR PIR; S00847; S00847.
 DR INTERPRO; IPR000495; -;
 DR INTERPRO; IPR003006; -;
 DR PFAM; PF00047; Ig; 3.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1

```
FT DOMAIN 1 97 CH1.  
FT DOMAIN 98 113 HINGE.  
FT DOMAIN 114 222 CH2.  
FT DOMAIN 223 329 CH3.  
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 143 203  
FT DISULFID 249 307  
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;  
  
Query Match 64.1%; Score 809.5; DB 1; Length 329;  
Best Local Similarity 63.8%; Pred. No. 1.5e-58;  
Matches 143; Conservative 41; Mismatches 37; Indels 3; Gaps 1;  
  
QY 12 PP---CPAPELLGGPSVLFPPPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 68  
DB 106 PPTDCSCDDNLGRPSVFIFPPKPKDITLTPKVTCTVVDVSEEPDQVQFSWFDVNR 165  
  
QY 69 VHNAKTRPREQYNSTYRVSVLVTHLDWLNKGYKCKVSNKALPAPIETISKAKGQP 128  
DB 166 VTAQTQPHSEQLNTRFVWSLTHIQHDMWSGKEFKCKVNNKDLPSPIETISKPRGKA 225  
  
QY 129 REPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 188  
DB 226 RTPQVYTTTPPREQMSKKNVSLTCNVTFYFASISVEVERNGELEDQKNTLPLVLDSDS 285  
  
QY 189 FFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 232  
DB 286 FYLSKLSVDTDMRGDIYCTCVVHEALHNHHTKQKLSRSPGK 329  
  
RESULT 14  
ID GCAB_MOUSE STANDARD; PRT; 330 AA.  
AC P01863;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE-81076554; PubMed-6777755;  
RA Sikorav J.-L., Auffray C., Rougeon F.;  
RT "Structure of the constant and 3' untranslated regions of the murine  
RT Balb/c gamma 2a heavy chain messenger RNA";  
RL Nucleic Acids Res. 8:3143-3155(1980).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-81198976; PubMed-6262729;  
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;  
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene  
RT and evolution of heavy chain genes: further evidence for intervening  
RT sequence-mediated domain transfer";  
RL Nucleic Acids Res. 9:1365-1381(1981).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-81223894; PubMed-6787604;  
RA Ollio R., Auffray C., Morchamps C., Rougeon F.;  
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes  
RT suggests that exons can be exchanged between genes in a multigenic  
RT family";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).  
RN [4]  
RP MYELOMA PROTEIN MOPC 173.  
RX MEDLINE-74175517; PubMed-4831970;  
RA Bourgois A., Fougereau M., Rocca-Serra J.;  
RT "Determination of the primary structure of a mouse IgG2a
```

```
RT immunoglobulin:amino-acid sequence of the Fc fragment. Implications  
RT for the evolution of immunoglobulin structure and function.";  
RL Eur. J. Biochem. 43:423-435(1974).  
RN [5]  
RX DISULFIDE BONDS.  
RX MEDLINE-73056887; PubMed-4565406;  
RA de Preval C., Fougereau M.;  
RT "Determination of the primary structure of a mouse gamma G2a  
RT immunoglobulin. Identification of the disulfide bridges";  
RL Eur. J. Biochem. 30:452-462(1972).  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; V00798; CAA24178.1; -  
DR PIR; A02152; G2MSA.  
DR INTERPRO; IPR000495; -  
DR INTERPRO; IPR003006; -  
DR PFAM; PF00047; Ig; 3.  
DR PROSITE; PS00290; IG_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON_TER 1 1 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 15 15  
FT DISULFID 27 82  
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.  
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;  
  
Query Match 64.1%; Score 809; DB 1; Length 330;  
Best Local Similarity 65.2%; Pred. No. 1.7e-58;  
Matches 146; Conservative 30; Mismatches 46; Indels 2; Gaps 1;  
  
QY 11 CPP---CPAPELLGGPSVLFPPPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 68  
DB 107 CPPCKPAPNLGGPSVFIFPPKIDVLMISLPIVTCVVVDSEDDPDVQISWFNVE 166  
  
QY 69 VHNAKTRPREQYNSTYRVSVLVTHLDWLNKGYKCKVSNKALPAPIETISKAKGQP 128  
DB 167 VHTAQQTQTHREDYNTSLRVSALPIQHDWMSGKEFKCKVNNKDLPSPIETISKPGSV 226  
  
QY 129 REPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 188  
DB 227 RAPQVYVLPPEEEMTKQVTLTCMVTFDMPEDIVETNNGKTELNTKPTPEVLDSDGS 286  
  
QY 189 FFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 232  
DB 287 FYMSKLRVEKKNVVERNSYSCSVVHEGLHNHHTKSFSTRPGK 330  
  
RESULT 15  
ID GCAB_MOUSE STANDARD; PRT; 335 AA.  
AC P01864;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2A CHAIN C REGION, B ALLELE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;
```

```

RX MEDLINE=82037861; PubMed=6170065;
RA Schreier P.H., Bothwell A.L.M., Mueller-Hill B., Baltimore D.;
RT "Multiple differences between the nucleic acid sequences of the
RL IgG2a and IgG2ab alleles of the mouse."
RL Proc. Natl. Acad. Sci. U.S.A. 78:4495-4499(1981).
RN [2]
RP SEQUENCE.
RX MEDLINE=82037777; PubMed=6794027;
RA Dognin M.J., Lauwereys M., Strosberg A.D.;
RT "Multiple amino acid substitutions between murine gamma 2a heavy
RT chain Fc regions of Ig1a and Ig1b allotypic forms."
RL Proc. Natl. Acad. Sci. U.S.A. 78:4031-4035(1981).
CC -!- MISCELLANEOUS: THE SEQUENCE DIFFERS FROM THAT OF THE A ALLELE,
CC FROM BALB/C MICE, AT 15% OF THE POSITIONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00479; -: NOT_ANNOTATED_CDS.
DR PIR: A02153; G2MSAB.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
SQ SEQUENCE 335 AA; 36596 MW; FA3382792CBB13C6 CRC64;
-----
Query Match 63.7%; Score 805; DB 1; Length 335;
Best Local Similarity 62.3%; Pred. No. 3.6e-58;
Matches 142; Conservative 38; Mismatches 48; Indels 0; Gaps 0;

QY 5 CDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 64
DB 108 CPHQVRVPCAPDPLGGPSVFIFPPKIKDVLMSLSPWVTCVVVDVSEDDPDVQISWV 167
QY 65 DGEVHNKATKPREQYNSTYRVSVLTFLHQDWLNGKEYKCKVSNKALPAPIEKTISK 124
DB 168 NNVEVHTAQTQTHREDYNSLTRVVSALPIQHQMWSGKEFKCKVNNRALSPIEKTISK 227
QY 125 KQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 184
DB 228 RGPVRAPOQVYVLPAPAEEMTKKEFSLTCMITGFLPAEIAVDWTSNGRTEQNKNTATVLD 287
QY 185 SDGSFFLYSKLTVDKSRWQGNVFCGSVMHEALHNHYTOKLSLSLSPCK 232
DB 288 SDGSYFWYSLRVQKSTWGRGSLFACSVHEVLNHLTTTKTISRSLGK 335
```

Search completed: March 1, 2001, 09:17:41
Job time: 259 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:13:22 ; Search time 79.26 seconds
(without alignments)
198.750 Million cell updates/sec

Title: US-09-389-782A-1
Perfect score: 1263
Sequence: 1 EPKSCDKTHRCPPCPAPELL.....MHEALHNYTKLSLSPGK 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1263	100.0	330	1 GHHU	Ig gamma-1 chain C
2	1257	98.5	374	2 S69339	Ig heavy chain V r
3	1257	99.5	374	2 S72664	Ig gamma-1 chain C
4	1255	99.4	255	4 S31866	Ig gamma-1 chain C
5	1210	95.8	234	2 PT0207	Ig gamma-3 chain C
6	1176	93.1	377	2 A23511	Ig gamma-3 chain C
7	1174	93.0	377	2 A60764	Ig gamma-3 chain C
8	1151	91.1	289	1 G3HUI	Ig gamma-3 heavy c
9	1145	90.7	326	1 G2HU	Ig gamma-2 chain C
10	1135	89.9	327	1 G4HU	Ig gamma-4 chain C
11	921	72.9	323	1 GHRB	Ig gamma chain C r
12	906.5	71.8	328	2 I47160	Ig gamma 2b chain
13	906.5	71.8	328	2 I47159	Ig gamma 2a chain
14	903	71.5	277	2 I47162	Ig gamma 4 chain c
15	896	70.9	329	1 G2GP	Ig gamma-2 chain c
16	885.5	70.1	328	2 I47158	Ig gamma 1 chain c
17	878.5	69.6	328	2 I47161	Ig gamma 3 chain c
18	856	67.8	470	2 S22080	Ig heavy chain pre
19	847.5	67.1	308	2 C30554	Ig heavy chain C r
20	847.5	67.1	472	2 S31459	Ig gamma-1 chain -
21	845.5	66.9	329	1 G3MSC	Ig gamma-3 chain C
22	842	66.7	333	2 PS0018	Ig gamma-2b chain
23	834.5	66.1	398	1 G3MSM	Ig gamma-3 chain C
24	832.5	65.9	444	2 PC4436	monoclonal antibody
25	822.5	65.1	324	1 G1MS	Ig gamma-1 chain C
26	822.5	65.1	326	2 PS0017	Ig gamma-1 chain C
27	817.5	64.7	393	1 G1MSM	Ig gamma-1 chain C
28	809.5	64.1	329	2 S00847	Ig gamma-2c chain
29	809	64.1	330	1 G2MSA	Ig gamma-2a chain

ALIGNMENTS

RESULT 1
GHHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999
C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:Z17370

A:Note: this sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(1) marker

A:Note: Lys-330 is removed after translation

R:Harris, L.J.

Submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <HAR>

A:Cross-references: EMBL:Z17370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of

A:Reference number: S33887; MUID:83001943

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113;235-330 <TAK>

A:Cross-references: EMBL:Z17370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edella

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq

A:Reference number: A90563; MUID:71064024

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96,'R',98-135 <GUN>

A:Note: this sequence has the Gln(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se

A:Reference number: A90564; MUID:71064025

A:Contents: Eu

A:Accession: A90564

A:Molecule type: protein

A:Residues: 136-154,'Q',156-165,'Q',178-194,'N',196-197,'D',199-238,'E',2

A:Note: this sequence has the Gln(non-1) markers, 239-Glu and 241-Met

R:Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A:Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni

Igen Primaerstruktur.

A;Reference number: A91668; MUID:77070269
 A;Contents: myeloma protein Nie
 A;Accession: B91668
 A;Molecule type: protein
 A;Residues: 1-34, Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27
 A;Note: this sequence has the Gln(17) and Gln(1) markers
 R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
 Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
 A;Title: Die Primaerstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1 KOI
 A;Reference number: A91723; MUID:83289131
 A;Contents: myeloma protein KOI; disulfide bonds
 A;Accession: A91723
 A;Molecule type: protein
 A;Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH
 A;Note: this sequence has the Gln(3) and Gln(non-1) markers
 R;Gall, W.E.; Edelman, G.M.
 Biochemistry 9, 3188-3196, 1970
 A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
 A;Reference number: A90565; MUID:71064027
 A;Contents: annotation; disulfide bonds
 R;Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
 Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
 A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
 enbromide cleavage products, and the disulfide bridges.
 A;Reference number: A91667; MUID:77070267
 A;Contents: annotation; disulfide bonds
 C;Genetics:
 A;Gene: GDB:IGHG1
 A;Cross-references: GDB:120085; OMIM:147100
 A;Map position: 14q32.33-14q32.33
 A;Introns: 99/1; 114/1; 224/1
 C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
 C;Superfamily: immunoglobulin C region; immunoglobulin homology
 C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F;20-85/Domain: immunoglobulin homology <IM1>
 F;137-206/Domain: immunoglobulin homology <IM2>
 F;243-310/Domain: immunoglobulin homology <IM3>
 F;27-83,144-204,250-308/Disulfide bonds: #status experimental
 F;103/Disulfide bonds: interchain (to light chain) #status experimental
 F;109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
 F;180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 1263; DB 1; Length 330;
 Best Local Similarity 100.0%; Pred. No. 8.4e-91;
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
 |||||
 Db 99 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 158
 QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLDQWLNGKEYCKCKVSNKALPAPIEKT 120
 |||||
 Db 159 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLDQWLNGKEYCKCKVSNKALPAPIEKT 218
 QY 121 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTP 180
 |||||
 Db 219 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTP 278
 QY 181 PVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232
 |||||
 Db 279 PVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 330

RESULT 2

S69339
 Ig heavy chain V region precursor - human
 C;Species: Homo sapiens (man)
 C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-2000
 C;Accession: S69339
 R;Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.

Eur. J. Biochem. 229, 54-60, 1995

A;Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
 A;Reference number: S69339; MUID:95262687
 A;Accession: S69339
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-374 <KHA>
 A;Cross-references: EMBL:X81695
 C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 99.5%; Score 1257; DB 2; Length 374;
 Best Local Similarity 99.1%; Pred. No. 2.9e-90;
 Matches 230; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
 |||||
 Db 143 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 202
 QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLDQWLNGKEYCKCKVSNKALPAPIEKT 120
 |||||
 Db 203 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLDQWLNGKEYCKCKVSNKALPAPIEKT 262
 QY 121 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTP 180
 |||||
 Db 263 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTP 322
 QY 181 PVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232
 |||||
 Db 323 PVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 374

RESULT 3

S72664
 Ig heavy chain V region precursor - human

C;Species: Homo sapiens (man)
 C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
 C;Accession: S72664
 R;Khamilichi, A.A.

submitted to the EMBL Data Library, September 1994

A;Reference number: S72664

A;Accession: S72664

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-374 <KHA>

A;Cross-references: EMBL:X81695

C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 99.5%; Score 1257; DB 2; Length 374;
 Best Local Similarity 99.1%; Pred. No. 2.9e-90;
 Matches 230; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
 |||||
 Db 143 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 202
 QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLDQWLNGKEYCKCKVSNKALPAPIEKT 120
 |||||
 Db 203 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLDQWLNGKEYCKCKVSNKALPAPIEKT 262
 QY 121 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTP 180
 |||||
 Db 263 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTP 322
 QY 181 PVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232
 |||||
 Db 323 PVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 374

RESULT 4

S31866
 Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000

C:Accession: S31866

R:Filipula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene products.

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FILL>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 99.4%; Score 1255; DB 4; Length 255;

Best Local Similarity 99.6%; Pred. No. 2.5e-90;

Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 24 EKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 83

QY 61 NWTVDGVEVHNATKPREQYNSTYRVSVTLVTLHODWLNKGEYKCKVSNKALPAPIEKT 120

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 84 NWTVDGVEVHNATKPREQYNSTYRVSVTLVTLHODWLNKGEYKCKVSNKALPAPIEKT 143

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 180

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 144 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 203

QY 181 PVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPGK 232

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 204 PVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPGK 255

RESULT 5

PT0207

Ig gamma chain C region - chimpanzee

C:Species: Pan troglodytes (chimpanzee)

C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999

C:Accession: PT0207

R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.

Mol. Immunol. 28, 319-322, 1991.

A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.

A:Reference number: PT0207; MUID:91287716

A:Accession: PT0207

A:Molecule type: mRNA

A:Residues: 1-234 <EHR>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 95.8%; Score 1210; DB 2; Length 234;

Best Local Similarity 98.7%; Pred. No. 7.1e-87;

Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 10 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 69

QY 61 NWTVDGVEVHNATKPREQYNSTYRVSVTLVTLHODWLNKGEYKCKVSNKALPAPIEKT 120

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 70 NWTVDGVEVHNATKPREQYNSTYRVSVTLVTLHODWLNKGEYKCKVSNKALPAPIEKT 129

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 180

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 130 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 189

QY 181 PVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPGK 225

Db 190 PVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPGK 234

RESULT 6

A23511

Ig gamma-3 chain C region (allotype G3m(b)) - human

C:Species: Homo sapiens (man)

C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999

C:Accession: A23511

R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.

Nucleic Acids Res. 14, 1779-1789, 1986

A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene:

A:Reference number: A23511; MUID:86148507

A:Accession: A23511

A:Molecule type: DNA

A:Residues: 1-377 <HUC>

A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056

C:Genetics:

A:Gene: GDB:IGHG3

A:Cross-references: GDB:119339; OMIM:147120

A:Map position: 14q32.33-14q32.33

A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 93.1%; Score 1176; DB 2; Length 377;

Best Local Similarity 92.7%; Pred. No. 5.6e-84;

Matches 215; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60

Db 146 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 205

QY 61 NWTVDGVEVHNATKPREQYNSTYRVSVTLVTLHODWLNKGEYKCKVSNKALPAPIEKT 120

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 206 NWTVDGVEVHNATKPREQYNSTYRVSVTLVTLHODWLNKGEYKCKVSNKALPAPIEKT 265

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 180

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 266 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 325

QY 181 PVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPGK 232

Db 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.||||| 1.|||||

Db 326 PVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPGK 377

RESULT 7

A60764

Ig gamma-3 chain C region, form LAT - human

C:Species: Homo sapiens (man)

C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999

C:Accession: A60764

R:Huck, S.; Lefranc, G.; Lefranc, M.P.

Immunogenetics 30, 250-257, 1989

A:Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conv

A:Reference number: A60764; MUID:90007613

A:Accession: A60764

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-377 <HUC>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 93.0%; Score 1174; DB 2; Length 377;

Best Local Similarity 92.7%; Pred. No. 8e-84;

Matches 215; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60

A:Reference number: A90253; MUID:72033500
 A:Contents: annotation; myeloma protein Sa, disulfide bonds
 R:Frangione, B.; Milstein, C.; Pink, J.R.L.
 Nature 221, 145-148, 1969
 A:Title: Structural studies of immunoglobulin G.
 A:Reference number: A93157; MUID:69064124
 A:Contents: annotation; Sa, disulfide bonds
 C:Genetics:
 A:Gene: GDB:IGHG2
 A:Cross-references: GDB:119338; OMIM:147110
 A:Map position: 14q32.33-14q32.33
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger complexes.
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F:20-85/Domain: immunoglobulin homology <IM1>
 F:133-202/Domain: immunoglobulin homology <IM2>
 F:239-306/Domain: immunoglobulin homology <IM3>
 F:14/Disulfide bonds: interchain (to light chain) #status experimental
 F:27-83,140-200,246-304/Disulfide bonds: #status experimental
 F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status predicted
 F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.7%; Score 1145; DB 1; Length 326;
 Best Local Similarity 91.4%; Pred. No. 1.2e-81;
 Matches 212; Conservative 9; Mismatches 7; Indels 4; Gaps 2;

QY 1 EPKSCDKTHCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
 DB 99 ERKCCVE---CPCPAPP-VAGSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQF 154
 QY 61 NWYVDGVEVHNKTPREQYNSTYRVSVLTIVLHODWLNKREYKCKVSNKALPAPIEKT 120
 DB 155 NWYVDGVEVHNKTPREQYNSTYRVSVLTIVLHODWLNKREYKCKVSNKALPAPIEKT 214
 QY 121 ISKAKQGPPEQYVYTPPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 180
 DB 215 ISKTKGQPREPQVYTPPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 274
 QY 181 PVLDSGSEFFLYSKLIVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232
 DB 275 PMLDSGSEFFLYSKLIVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 326

RESULT 10
 G4HU
 Ig gamma-4 chain C region - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999
 C:Accession: A90933; A90249; A02150
 R:Ellison, J.; Buxbaum, J.; Hood, L.
 DNA 1, 11-18, 1981
 A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
 A:Reference number: A90933; MUID:83157104
 A:Accession: A90933
 A:Molecule type: DNA
 A:Residues: 1-327 <ELL>
 A:Note: the sequence was determined from the germline gene
 R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
 Biochem. J. 117, 33-47, 1970
 A:Title: Human immunoglobulin subclones. Partial amino acid sequence of the constant region.
 A:Reference number: A90249; MUID:70207560
 A:Accession: A90249
 A:Molecule type: protein
 A:Residues: 1-30;81-326 <PIN>
 C:Genetics:
 A:Gene: GDB:IGHG4
 A:Cross-references: GDB:119340; OMIM:147130
 A:Map position: 14q32.33-14q32.33
 A:Introns: 99/1; 11q/1; 22/1
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger complexes.

C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F:20-85/Domain: immunoglobulin homology <IM1>
 F:99-110/Region: hinge
 F:134-203/Domain: immunoglobulin homology <IM2>
 F:240-307/Domain: immunoglobulin homology <IM3>
 F:14/Disulfide bonds: interchain (to light chain) #status experimental
 F:27-83,141-201,247-305/Disulfide bonds: #status predicted
 F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
 F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.9%; Score 1135; DB 1; Length 327;
 Best Local Similarity 93.7%; Pred. No. 7.1e-81;
 Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 11 CPCCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 70
 DB 106 CPSCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVH 165
 QY 71 NAKTKPREQYNSTYRVSVLTIVLHODWLNKREYKCKVSNKALPAPIEKTISKAKGPPE 130
 DB 166 NAKTKPREQYNSTYRVSVLTIVLHODWLNKREYKCKVSNKALPAPIEKTISKAKGPPE 225
 QY 131 PQVYTPPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSEFF 190
 DB 226 PQVYTPPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSEFF 285
 QY 191 LYSKLVIVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232
 DB 286 LYSRLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 327

RESULT 11
 CHRB
 Ig gamma chain C region - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
 C:Accession: A91749; A90290; A93928; A90245; A94416; A02161
 R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
 Immunogenetics 18, 387-397, 1983
 A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I hap
 A:Reference number: A91749; MUID:84030930
 A:Accession: A91749
 A:Molecule type: mRNA
 A:Residues: 1-323 <BER>
 A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-
 R:Pratt, D.M.; Mole, L.E.
 Biochem. J. 151, 337-349, 1975
 A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunog
 A:Reference number: A90290; MUID:76135469
 A:Accession: A90290
 A:Molecule type: protein
 A:Residues: 1-47,'E',49-71,'PV',72-128 <PRA>
 R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
 Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
 A:Title: Heavy chain genes of rabbit IgG: Isolation of a cDNA encoding gamma heavy ch
 A:Reference number: A93928; MUID:83299917
 A:Accession: A93928
 A:Molecule type: mRNA
 A:Residues: 88-103,'M',105-143,'E',145-184,'A',186,'E',188-266 <MAR>
 A:Cross-references: GB:M16426; NID:g165111; PIDN:AAA31289.1; PID:g165112
 A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic ma
 R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
 Biochem. J. 116, 249-259, 1970
 A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul
 A:Reference number: A90245; MUID:70110015
 A:Accession: A90245
 A:Molecule type: protein
 A:Residues: 132-143,'E',145-161 <FRU>
 R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
 In Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wikse
 A:Reference number: A94416

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:41 ; Search time 40.97 Seconds
(without alignments)
316.083 Million cell updates/sec

Title: US-09-389-782A-2
Perfect score: 2179
Sequence: 1 MNRWLCCALLVLDIIEWTT.....QKLFLEMIGNQVQSVKISCL 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	314	14.4	461	1	TNR2_HUMAN
2	312	14.3	474	1	TNR2_MOUSE
3	270	12.4	289	1	CD40_MOUSE
4	261	12.0	277	1	CD40_HUMAN
5	244	11.2	435	1	TNRC_HUMAN
6	233.5	10.7	325	1	VT2_SFVKA
7	228	10.5	269	1	CD40_BOVIN
8	220	10.1	326	1	VT2_MXKVL
9	217	10.0	349	1	VC22_VARV
10	206	9.5	415	1	TNRC_MOUSE
11	199	9.1	427	1	NGFR_HUMAN
12	197.5	9.1	425	1	NGFR_RAT
13	190.5	8.7	323	1	FASA_BOVIN
14	184	8.4	416	1	NGFR_CHICK
15	182	8.4	335	1	FASA_HUMAN
16	177	8.1	256	1	41BB_MOUSE
17	176	8.1	332	1	FASA_PIG
18	167	7.7	255	1	41BB_HUMAN
19	165.5	7.6	461	1	TNR1_PIG
20	163	7.5	271	1	OX40_RAT
21	161.5	7.4	454	1	TNR1_MOUSE
22	160	7.3	455	1	TNR1_HUMAN
23	160	7.3	471	1	TNR1_BOVIN
24	153	7.0	461	1	TNR1_RAT
25	152.5	7.0	272	1	OX40_MOUSE
26	151	6.9	277	1	OX40_HUMAN
27	147.5	6.8	324	1	FASA_RAT
28	143	6.6	327	1	FASA_MOUSE
29	139	6.4	595	1	CD30_HUMAN
30	133	6.1	1786	1	LMB1_HUMAN
31	128.5	5.9	1790	1	LMB1_MOUSE
32	126.5	5.8	260	1	CD27_HUMAN
33	125	5.7	1680	1	FUR2_HUMAN
					P30432 drosophila

34 120.5 5.5 250 1 CD27_MOUSE
35 120.5 5.5 2911 1 FBN2_HUMAN
36 119.5 5.5 1639 1 LMGI_DROME
37 119 5.5 1609 1 LMGI_HUMAN
38 118 5.4 3712 1 LMA_DROME
39 116.5 5.3 915 1 PAC6_MOUSE
40 116 5.3 1955 1 AGRI_CHICK
41 116 5.3 3084 1 LMA1_MOUSE
42 116 5.3 3106 1 LMA2_MOUSE
43 115.5 5.3 713 1 TSA4_GIALA
44 115.5 5.3 1607 1 LMGI_MOUSE
45 115 5.3 417 1 WSL1_HUMAN

ALIGNMENTS

RESULT 1
ID TNR2_HUMAN STANDARD; PRT; 461 AA.
AC P20333;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR
DE BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).
GN TNFRSF1B OR TNFR2 OR TNFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90260639; PubMed=2160731;
RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,
RA Dower S.K., Cosman D., Goodwin R.G.;
RT "A receptor for tumor necrosis factor defines an unusual family of
RT cellular and viral proteins.";
RL Science 248:1019-1023(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91045991; PubMed=2172983;
RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,
RA Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;
RT "A second tumor necrosis factor receptor gene product can shed a
RT naturally occurring tumor necrosis factor inhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96299745; PubMed=8661109;
RA Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,
RA Lepaslier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,
RA Brodeur G.M.;
RT "Physical mapping and genomic structure of the human TNFR2 gene.";
RL Genomics 35:94-100(1996).
RN [4]
RP SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90349572; PubMed=2166946;
RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,
RA Ringold G.M.;
RT "Complementary DNA cloning of a receptor for tumor necrosis factor
RT and demonstration of a shed form of the receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
RN [5]
RP SEQUENCE OF 27-31.
RX MEDLINE=90110215; PubMed=2153136;
RA Engelmann H., Novick D., Wallach D.;
RT "Two tumor necrosis factor-binding proteins purified from human
RT urine. Evidence for immunological cross-reactivity with cell surface
RT tumor necrosis factor receptors.";
RL J. Biol. Chem. 265:1531-1536(1990).
RN [6]
RP SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.
RX MEDLINE=91056048; PubMed=2173696;

P41272 mus musculus
P35556 homo sapien
P15215 drosophila
P11047 homo sapien
Q00174 drosophila
Q04592 mus musculus
P31696 gallus gall
P19137 mus musculus
Q06675 mus musculus
P21849 giardia lam
P02468 mus musculus
Q93038 h wsl-1 pro

*Molecular cloning and expression of the type 1 and type 2 murine
RT receptors for tumor necrosis factor.*;
RT Mol. Cell. Biol. 11:3020-3026(1991).
RN [3]
RN SEQUENCE OF 1-26 FROM N.A.
RC STRAIN=NOD;
RA Jacob C.O., Liu J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE OF 1-22 FROM N.A.
RC TISSUE=LIVER;
RA Kissingerhis M., Fellores R., Feldmann M., Chernaiovsky Y.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M50469; AAA39752.1; -
DR EMBL; M59378; AAA40463.1; -
DR EMBL; U39488; AAA85021.1; -
DR EMBL; X87128; CAA60618.1; -
DR PIR; B38634; B38634.
DR HSSP; P19438; INCF.
DR MGD; MGI:1314883; TNFRSF1B.
DR INTERPRO: IPR001368; -
DR PFAM: PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
DR Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 474 TUMOR NECROSIS FACTOR RECEPTOR 2.
FT DOMAIN 23 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 259 288 POTENTIAL.
FT DOMAIN 289 474 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 39 203 4 X TNFR-CYS.
FT REPEAT 39 77 TNFR-CYS 1.
FT REPEAT 78 119 TNFR-CYS 2.
FT REPEAT 120 164 TNFR-CYS 3.
FT REPEAT 165 203 TNFR-CYS 4.
FT DISULFID 40 54 BY SIMILARITY.
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 59 76 BY SIMILARITY.
FT DISULFID 79 94 BY SIMILARITY.
FT DISULFID 97 111 BY SIMILARITY.
FT DISULFID 101 119 BY SIMILARITY.
FT DISULFID 121 127 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 139 163 BY SIMILARITY.
FT DISULFID 166 181 BY SIMILARITY.
FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 474 AA; 50319 MW; 462EAE398C4D6563 CRC64;

Query Match 14.3%; Score 312; DB 1; Length 474;
Best Local Similarity 34.3%; Pred. No. 2.1e-15;
Matches 70; Conservative 22; Mismatches 84; Indels 28; Gaps 5;
5 LCCALLVLLDIIEWTQETLPKYL--HYDPETGH-----QLLCDKCAPGT 49
6 LWALVFELQL--WATCHTPPAQVLLTPYKPEPGYEQISQEYIDRAQMCCKAPGQY 63
50 LKQCHTVRRKTLVCPDPDHSWHTSDVCYVCSPVKELQSKVQBCNTHNRVCSEE 109
64 VKHFCNKTSITVCDACEASMTQWNQFRCLSSSCSTTDQVEIRACTKQNRVCAEEA 123

110 GRYLEIEF-----CLKHRSPPGSGVVOAGTPERNTVCKKCPDGFSGFTSSKAPCIK 162
124 GRYCALKTHSGSCRCQRLSKCGFGVASSRAPNGVNLKACAPGTFSDTTS2DVCRP 183
163 HTNCSTFGLLLIQGNATHDNVCS 186
184 HRICS-----ILAIPGNASTDAVCA 203
RESULT 3
CD40_MOUSE STANDARD; PRT; 289 AA.
ID CD40_MOUSE AC P27512;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40).
GN TNFRSF5 OR CD40.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92105763; PubMed-1370315;
RA Torres R.M., Clark E.A.;
RT "Differential increase of an alternatively polyadenylated mRNA
species of murine CD40 upon B lymphocyte activation.";
RL J. Immunol. 148:620-626(1992).
RN [2]
RP REVISIONS.
RC STRAIN=BALB/C;
RA Torres R.M.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=LIVER;
RX MEDLINE-93094586; PubMed-1281194;
RA Grimaldi J.C., Torres R., Kozak C.A., Chang R., Clark E.A.,
RA Howard M., Cockayne D.A.;
RT "Genomic structure and chromosomal mapping of the murine CD40 gene.";
RL J. Immunol. 149:3921-3926(1992).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M83312; AAB08705.1; -
DR EMBL; M94126; AAA37404.1; -
DR EMBL; M94129; AAA37404.1; JOINED.
DR EMBL; M94128; AAA37404.1; JOINED.
DR EMBL; M94127; AAA37404.1; JOINED.
DR PIR; A46476; A46476.
DR HSSP; P25942; 1CDF.
DR MGD; MGI:88336; TNFRSF5.
DR INTERPRO: IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS50050; TNFR_NGFR_2; 4.
KW Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 289 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 289 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.

```

FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 289 AA; 32111 MW; C791CB6D2FEA574E CRC64;

Query Match 12.4%; Score 270; DB 1; Length 289;
Best Local Similarity 34.5%; Pred. No. 1.2e-12;
Matches 57; Conservative 26; Mismatches 58; Indels 14; Gaps 5;

QY 27 KYLHYDPETHQLKDKCAPGTYLKQHCVRKTLVCPDHSYTDSWHTSDEC---VYC 83
DB 30 QYLH-DGQ-----CDLCPGSRSLTSHCTALEKTCHEPCDSGEFSAQNNREIRCHOHRHC 83

QY 84 SPVKELQSVKQECNRTHNRVCECEGRYL---EIEFLKHKRSCPPGSGVQAGTPERT 140
DB 84 EP--NQGLRVKKGTAESDVTCTCKEGHCTSKDCEACQAHTPCIPGFGVMEMATETTDT 141

QY 141 VCKKCPDGFSSGTSKAPCIKHTNCSTFGLLIIOGNATHDNVC 185
DB 142 VCHPCPVGFFSNQSLFEKCPWTSCEKDNLEVLQKGTSTQNVIC 186

RESULT 4
CD40_HUMAN STANDARD; PRT; 277 AA.
AC P25942;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40)
DE (TUMOR NECROSIS FACTOR RECEPTOR 5).
GN TNFRSF5 OR CD40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89355608; PubMed=2475341;
RA Stamenkovic I., Clark E.A., Seed B.;
RT "A B-lymphocyte activation molecule related to the nerve growth
RT factor receptor and induced by cytokines in carcinomas."
RL EMBO J. 8:1403-1410(1989).
RN [2]
RP 3D-STRUCTURE MODELING OF 24-144.
RX MEDLINE=97189482; PubMed=9037712;
RA Bajorath J., Aruffo A.;
RT "Construction and analysis of a detailed three-dimensional model of
RT the ligand binding domain of the human B cell receptor CD40L."
RL Proteins 27:59-70(1997).
RN [3]
RP 3D-STRUCTURE MODELING OF 26-186 IN COMPLEX WITH CD40L.
RX MEDLINE=98266353; PubMed=9605317;
RA Singh J., Garber E., van Vlijmen H., Karpsusas M., Hsu Y.-M.,
RA Zheng Z., Naismith J.H., Thomas D.;
RT "The role of polar interactions in the molecular recognition of CD40L
RT with its receptor CD40."
RL Protein Sci. 7:1124-1135(1998).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD40 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd40.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60592; CAA43045.1;
DR PIR; S04460; S04460
DR PDB; ICDF; 01-APR-97.
DR MIM; 109535;
DR INTERPRO; IPR001368;
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS00500; TNFR_NGFR_2; 4.
KW Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal;
KW 3D-structure.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 277 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 61 103 TNFR-CYS 1.
FT REPEAT 104 144 TNFR-CYS 2.
FT REPEAT 145 187 TNFR-CYS 3.
FT DISULFID 26 37 TNFR-CYS 4.
FT DISULFID 38 51
FT DISULFID 41 59
FT DISULFID 62 77
FT DISULFID 83 103
FT DISULFID 105 119
FT DISULFID 111 116
FT DISULFID 125 143
FT CARBOHYD 153 153
FT CARBOHYD 180 180
SQ SEQUENCE 277 AA; 30619 MW; BC8776EC2C4A5680 CRC64;

Query Match 12.0%; Score 261; DB 1; Length 277;
Best Local Similarity 33.9%; Pred. No. 5.1e-12;
Matches 58; Conservative 23; Mismatches 74; Indels 16; Gaps

QY 29 LHYDPETHG-----HQLLCKCAPGTYLKQHCVRKTLVCPDHSYTDSWHTSDEC 80
DB 18 VHPPTACREKQYLNQSCSLCPGKQKLVSDCTETETELCPGSEFELDTWNRTHC 77

QY 81 ---VYCSVPVKELQSVKQECNRTHNRVCECEGRYL---EIEFLKHKRSCPPGSGVQAG 134
DB 78 HQHKYCDNLG--LRVQKGTSETDTICTCEGHCTSEACESCVLHRSCTPGFGVKQIA 135

QY 135 TPERNTVCKKCPDGFSSGTSKAPCIKHTNCSTFGLLIIOGNATHDNVC 185
DB 136 TGVSDTICEPCPVGFFSNVSSAFKCHPWTSCETKDLVVQAGTKNTDVC 186

RESULT 5
TNRC_HUMAN STANDARD; PRT; 435 AA.
AC P36941;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LYPHOTOMIN-BETA RECEPTOR (TUMOR NECROSIS FACTOR RECEPTOR
DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
GN LTR OR TNFR OR TNFRSF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LIVER;
RX MEDLINE=93252381; PubMed=8486360;
RA Beens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
RT "Construction and evaluation of a hncDNA library of human l2p
RT transcribed sequences derived from a somatic cell hybrid."
RL Genomics 16:214-218(1993).

```

RN FUNCTION.
 RP MEDLINE=94225209; PubMed=8171323;
 RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,
 RA Ehrenfeld B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
 RT "A lymphotoxin-beta-specific receptor.";
 RL Science 264:707-710(1994).
 CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
 CC IMMUNE DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L04270; AAA36757.1; -;
 DR HSP; P25942; 1CDF.
 DR MIM; 600979;
 DR INTERPRO; IPR001368; -;
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00050; TNFR_NGFR_2; 3.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 435 LYMPHOTOXIN-BETA RECEPTOR.
 FT DOMAIN 31 227 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 228 248 POTENTIAL.
 FT DOMAIN 249 435 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 42 211 4 X TNFR-CYS.
 FT REPEAT 42 81 TNFR-CYS 1.
 FT REPEAT 82 124 TNFR-CYS 2.
 FT REPEAT 125 168 TNFR-CYS 3.
 FT REPEAT 169 211 TNFR-CYS 4.
 FT DISULFID 43 58 BY SIMILARITY.
 FT DISULFID 59 72 BY SIMILARITY.
 FT DISULFID 62 80 BY SIMILARITY.
 FT DISULFID 83 98 BY SIMILARITY.
 FT DISULFID 101 116 BY SIMILARITY.
 FT DISULFID 104 124 BY SIMILARITY.
 FT DISULFID 126 132 BY SIMILARITY.
 FT DISULFID 139 148 BY SIMILARITY.
 FT DISULFID 142 167 BY SIMILARITY.
 FT DISULFID 170 185 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 435 AA; 46709 MW; 624626B5022F656F CRC64;
 QY 9 LVLVLLDIETWTQETLPP-----KYLHYDETGHQLLCKDCAPGTYLKQHCVTVR 57
 DB 18 VLGLFGLLAASQPAQVPPYASENQTCRDOKEYEYEPQ--HRIICSCRPPTGYYSKCSRI 75
 QY 58 RKTLCVCPDHSHTSDWHTSDEVCYSPCKELQSKQ--ECNRTHNRVCECEGRY--- 112
 DB 76 RDTVCATCAENSNEHWNLYITICQLCRP-CDPVNGEEIAPCTSKRKTQCRQGMFCAA 134
 QY 113 --LEIEFLKHSRCPGSG--GVVQAGTPERTNVCKPDGFFSGTSSKAPCIKHTNCSTF 169
 DB 135 WALECTHCELLSDCPGTEAELEKDEVGKGNHCVCKAGHFQNTSSPSARCPHTRCENQ 194
 QY 170 GLLLIKNATHNVCNSGNE 190
 DB 195 GLVEAAPGTAQSDTTCKNPLE 215

RESULT 6
 VT2_SFVKA STANDARD; PRT; 325 AA.
 AC P25943;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
 GN T2.
 OS Shope fibroma virus (strain Kasza) (SFV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Leporipoxvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87321103; PubMed=2820128;
 RA Upton C., Delange A.M., McFadden G.;
 RT "Tumorigenic poxviruses: genomic organization and DNA sequence of the
 RT telomeric region of the Shope fibroma virus genome.";
 RL Virology 160:20-30(1987).
 RN [2]
 RP FUNCTION
 RX MEDLINE=91207415; PubMed=1850261;
 RA Smith C.A., Davis T., Wignall J.M., Din W.S., Farrah T., Upton C.,
 RA McFadden G., Goodwin R.G.;
 RT "T2 open reading frame from the Shope fibroma virus encodes a soluble
 RT form of the TNF receptor.";
 RL Biochem. Biophys. Res. Commun. 176:335-342(1991).
 CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
 CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
 CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M17433; -; NOT_ANNOTATED_CDS.
 DR EMBL; A23727; CAA01687.1; -;
 DR PIR; B43692; B43692.
 DR HSP; P19438; 1TNR.
 DR INTERPRO; IPR001368; -;
 DR PFAM; PF00020; TNFR_C6; 2.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00050; TNFR_NGFR_2; 1.
 KW Receptor; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 325 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
 FT DOMAIN 27 186 4 X TNFR-CYS.
 FT REPEAT 27 62 TNFR-CYS 1.
 FT REPEAT 63 104 TNFR-CYS 2.
 FT REPEAT 105 147 TNFR-CYS 3.
 FT REPEAT 148 186 TNFR-CYS 4.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 325 AA; 35132 MW; 810530339198A71E CRC64;
 QY 36 GHQ-----LLCDKCAPGTYLKQHCVTVRKTLVCPDHSHTSDWHTSDEVCYSPCKEL 90
 DB 30 GHDEYKDLGLCCASCHPGFYASRLCGPSGNTVCSPECDGTFTASTNHAPACVSCRGCTGH 89
 QY 91 QSVKQECNRTHNRVCECEGRYLEIE-----FCLKHRSPPGSGVGVQAGTPERTNVCKK 144
 Query Match 10.7%; Score 233.5; DB 1; Length 325;
 Best Local Similarity 31.7%; Pred. No. 5.8e-10;
 Matches 59; Conservative 19; Mismatches 91; Indels 17; Gaps 5;

```

Db 90 LSESPQCDRTDRVNCNCSTGNYCLLKQNGORICAPQTKCPAGYV-VSGHTRAGDTLCEK 148
QY 145 CPDGFSGSETSSKAPCIKHTNCSFGLLLIOKNATHDNVCSGNREATAKCGIDVTLL--- 201
Db 149 CPPIHTYDLSPTBRCGTSFNYISVGFNLYPV-NETSCTTTAGHNEVIKTEFTVLTNYT 207
QY 202 -CEEAF 206
Db 208 DCDPVF 213

RESULT 7
CD40_BOVIN STANDARD; PRT; 269 AA.
AC Q28203;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (FRAGMENT).
GN TNFRSF5 OR CD40.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=97281252; PubMed=9135560;
RA Hirano A., Brown W.C., Estes D.M.;
RT "Cloning, expression and biological function of the bovine CD40
RT homologue: role in B-lymphocyte growth and differentiation in
RT cattle.";
RL Immunology 90:294-300(1997).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U57745; AAC48710.1;
CC HSP; P25942; ICDF
CC INTERPRO; IPR001368;
CC PFAM; PF00020; TNFR_C6; 4.
CC PROSITE; PS00652; TNFR_NGFR_1; 1.
CC PROSITE; PS50050; TNFR_NGFR_2; 1.
CC Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 >269 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215
FT DOMAIN 216 >269 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 216 187 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 269 269
SQ SEQUENCE 269 AA; 29983 MW; 746903F30F95F387 CRC64;

Query Match 10.5%; Score 228; DB 1; Length 269;
Best Local Similarity 31.9%; Pred. No. 1.le-09;
Matches 60; Conservative
QY 29 LHYDPETG-----HOLLCDKCAPGTLYLKQHCTVRRKTLVCPDPHSYDSWHTSDEC 80

```

```

Db 18 VHSEPATACGEKQYPVNSLCCDLPPGQKLVNDCTEVSKTEQSCGKGFEFLSTNRKVC 77
QY 81 ---VYCSP-VCKELOSQVQBCNRRHNRVCEBGRYL---EIFELCKHRSPPGSGVQVA 133
Db 78 HEHRYCNFNLAGRIQS---EGTLNTDTTCVCEGQHCSTHTCESCTPHSLCLUPFGVAKI 134
QY 134 GTPERTNTVCKKCPDGFSGSETSSKAPCIKHTNCSFGLLLIOKNATHDNVCSGNREATAQ 193
Db 135 ATGLLDTVCECPGLGFFSNVSAFEKCHRWTSCEKRGVLVEQHVGNKNTDVCVGGESRMT 194
QY 194 KCGIDVTLL 201
Db 195 LWIPVTM 202

RESULT 8
VT2_MYXVL STANDARD; PRT; 326 AA.
AC P29825;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN 72).
GN TN2.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=91335768; PubMed=1651597;
RA Upton C., Macen J.L., Schreiber M., McFadden G.;
RT "Myxoma virus expresses a secreted protein with homology to the tumor
RT necrosis factor receptor gene family that contributes to viral
RT virulence.";
RL Virology 184:370-382(1991).
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M95181; AAA46632.1;
CC EMBL; A23729; CAA01688.1;
CC FIR; A40566; GQVZML.
CC HSP; P19438; ITNR.
CC INTERPRO; IPR001368;
CC PFAM; PF00020; TNFR_C6; 2.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC PROSITE; PS50050; TNFR_NGFR_2; 2.
CC Receptor; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 326 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
FT DOMAIN 27 186 4 X TNFR-CYS.
FT REPEAT 27 62 TNFR-CYS 1.
FT REPEAT 63 104 TNFR-CYS 2.
FT REPEAT 105 147 TNFR-CYS 3.
FT REPEAT 148 186 TNFR-CYS 4.
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 326 AA; 35208 MW; ABBF027E947292FF CRC64;

Query Match 10.1%; Score 220; DB 1; Length 326;

```

Gaps

Gaps

Gaps

Gaps

Gaps

Gaps

Db	11	FLSCIIINGDAAPYT-----PNGKCKDTEYKRHNLCCLSCPDPGTYASRLCDSKNTQOC	65
Qy	63	VPCDPDHSYTDSWHTSDCYVCSPVKELQSVKQECNRTHNRVCECEBEGRYLEI-----E	116
Db	66	TPCGSGFTSRNNHLPACLSGNGRCNSNOVETRSCTTHNRICECSPGYYCLLKSGSGCK	125
Qy	117	FCLKHRSCPPGSGVVGAGTPEPRTVNTVKKCPDGFPSGETSSKAPCIKHTNCSTFGLLLIQ-	175
Db	126	ACVSQTKCGIGYG-VSGHTSWGDVICSPCGFGTYSHYTVSSADCKECPVPN-NTENYIDVEI	183
Qy	176	KGNATHDNVCSGNREATQKCGIDVTLCEEAFFFAVPYTKIIPNWLSVLVDLSLPGTKVNAE	235
Db	184	TLYPVNDTSTCT---RTTTLGSE-----SILTSEL-----	210
Qy	236	SVERIKRRHSS-----QEOTFOLLK-----LWKHONRDOEMVKKIIQDIDLCESSVORH	284
Db	211	---TITWNHDDCPVFRFEYFVSLNKVATSGFTGENRYQNIK-----VCTLNFEIK	260
Qy	285	LGHSNLTTEOLLALMESLPKKISPEEIERTRTKCSSEOLLKLLSLWRINKGDDQDTLK	344
Db	261	CNNKSGSFQOLTRAKND--DGMASHSETVTLAGDCLSSVDIYILSYNTNAQDYETDTS-	317
Qy	345	LMYALKHL--KTHFP-----KTVTHSLRKTWRFL	372
Db	318	--YRVGNVLDDSHMPGSCNIHKPI TNS--KPTRFL	349
RESULT 10.			
ID	TNRC_MOUSE	STANDARD;	PRT; 415 AA.
ID	TNRC_MOUSE		

RESULT

RESULT

Query Match

Best Local s

Matches 92

4 WLCC
:| |

Qy	149	FFSGEYTSKAPCIKHKNCTFTGLLLI-----QKGNATHD--NYVCSGNREAT	199
Db	164	--GSR-----HAN-SLWALLILLPIVLIIYKVKSRNKKNDYCSAASNDEGR	212
Qy	193	QKCGIDVTLCEEAFFFAVPTKIIPNWSLVLDLSPGTK-----VNAESVERIK-----	241
Db	213	QLNLTDVDL-----GKYPS-----IAEQMRITEVKFVRKNGMEEAKIDDIIMHD	257
Qy	242	RRISSOEQTFOILLKLW--KHONRD-----QEMVKKIIODI-DLCESSVORHLGHSNLT	292
Db	258	NVHETAEQKVOLLRNWYQSHGKNAYCTLTKSLPKALAEKIDYMKDITNERENANLON	317
Qy	293	E 293	
Db	318	E 318	
RESULT 14			
NGFR_CHK			
ID	NGFR_CHK	STANDARD; PRT; 416 AA.	
AC	PI8519;		
DT	01-NOV-1990 (Rel. 16, Created)		
DT	01-NOV-1990 (Rel. 16, Last sequence update)		
DT	01-OCT-2000 (Rel. 40, Last annotation update)		
DE	LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)		
DE	(GP80-LNGFR) (P75 ICD).		
GN	NGFR		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
RC	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RC	MEDLINE=90166579; PubMed=2560385;		
RA	Large T.H., Weskamp G., Helder J.C., Radeke M.J., Misko T.P.,		
RA	Shooter E.M., Reichardt L.F.;		
RT	"Structure and developmental expression of the nerve growth factor		
RT	receptor in the chicken central nervous system.";		
RL	Neuron 2:1123-1134(1989).		
RN	[2]		
RC	SEQUENCE OF 21-416 FROM N.A.		
RA	MEDLINE=90152140; PubMed=2154393;		
RA	Heuer J.G., Fatemie-Nainie S., Wheeler E.F., Bothwell M.;		
RT	"Structure and developmental expression of the chicken NGF receptor."		
RT	Dev. Biol. 137:287-304(1990).		
CC	-1- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,		
CC	NT-3, AND NT-4.		
CC	-1- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE		
CC	BOND FORMATION.		
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.		
CC	-1- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.		
CC	-1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.		
CC	-1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.		
DR	PIR; JN0006; JN0006.		
DR	PIR; A60504; A60504.		
DR	HSP; P07174; INGR.		
DR	INTERPRO; IPR000488; .		
DR	INTERPRO; IPR001368; .		
DR	PFAM; PF00020; TNFR_c6; 4.		
DR	PFAM; PF00531; death; 1.		
DR	PROSITE; PS00652; TNFR_NGFR.1; 3.		
DR	PROSITE; PS00050; TNFR_NGFR.2; 3.		
DR	PROSITE; PS50017; DEATH_DOMAIN; 1.		
KW	Receptor; Neurogenesis; Transmembrane; Glycoprotein; Repeat;		
KW	Phosphorylation; Signal.		
FT	SIGNAL	1 19	
FT	CHAIN	20 416	
FT	DOMAIN	29 239	
FT	TRANSMEM	240 261	
FT	DOMAIN	262 416	
FT	DOMAIN	23 181	

```

FT REPEAT 23 57 TNFR-CYS 1.
FT REPEAT 58 100 TNFR-CYS 2.
FT REPEAT 101 140 TNFR-CYS 3.
FT REPEAT 141 181 TNFR-CYS 4.
FT DOMAIN 188 236 SER/THR-RICH.
FT DOMAIN 333 410 DEATH DOMAIN.
FT DISULFID 24 35 BY SIMILARITY.
FT DISULFID 36 49 BY SIMILARITY.
FT DISULFID 39 56 BY SIMILARITY.
FT DISULFID 59 75 BY SIMILARITY.
FT DISULFID 78 91 BY SIMILARITY.
FT DISULFID 81 99 BY SIMILARITY.
FT DISULFID 101 114 BY SIMILARITY.
FT DISULFID 117 130 BY SIMILARITY.
FT DISULFID 120 138 BY SIMILARITY.
FT DISULFID 141 156 BY SIMILARITY.
FT DISULFID 159 172 BY SIMILARITY.
FT DISULFID 162 180 BY SIMILARITY.
FT CARBOHYD 52 52 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 36 36 C -> Y (IN REF. 2).
FT CONFLICT 173 173 T -> K (IN REF. 2).
FT CONFLICT 276 276 N -> S (IN REF. 2).
FT CONFLICT 396 396 K -> R (IN REF. 2).
SQ SEQUENCE 416 AA; 44654 MW; 6BCEAB54F4D2D56 CRC64;

Query Match 8.4%; Score 184; DB 1; Length 416;
Best Local Similarity 30.9%; Pred. No. 2.7e-06;
Matches 46; Conservative 28; Mismatches 67; Indels 8; Gaps 6;

QY 41 CDKCAPGTYLKHQVRRKTLVCPDPH-SYDTSWHTSDCVYKSPVKELQSVKQECNR 99
DB 36 CKACNLGEGVVPQGV-NQTVCEPCLDSVTYSVATEPCPKCTQ-CVGLHSMSPCVE 93
QY 100 THNRVCECEGRYLETEF---CLKHRSCPPGSGVWQAGTPERTVCKKCPDGFSGTSS 156
DB 94 SDDAVCRCAYG-YFQDELSSCKECSCEVGFGLMFPFCDSDQTVCECEPGETFSDEANF 152
QY 157 KAPCIKHTNGSTFGLLLIOGNATHDNVC 185
DB 153 VDCPLCTICEE-NEVMVKCTATSDAEC 180

RESULT 15
FASA_HUMAN STANDARD; PRT; 335 AA.
AC P25445;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95 ANTIGEN).
GN TNFRSF6 OR APT1 OR FAS OR FASL.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE-91309137; PubMed-1713127;
RA Itom N., Yonehara S., Ishii A., Yonehara M., Mizushima S.I.,
RA Sameshima M., Hase A., Seto Y., Nagata S.;
RT "The polypeptide encoded by the cDNA for human cell surface antigen
RL Fas can mediate apoptosis.";
RL Cell 66:233-243(1991).
[2]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 226-240; 269-291 AND 321-335.
RX MEDLINE-92268122; PubMed-1375228;
RX Oehm A., Behrmann I., Falk W., Pawlita M., Maier G., Klas C.,
RA Li-Weber M., Richards S., Dhein J., Trauth B.C., Ponstingl H.,
RA Kramer P.H.;
RT "Purification and molecular cloning of the APO-1 cell surface
RT antigen, a member of the tumor necrosis factor/nerve growth factor
RT receptor superfamily. Sequence identity with the Fas antigen.";
```

```

J. Biol. Chem. 267:10709-10715(1992).
[3]
RN STRUCTURE BY NMR OF 218-335.
RX MEDLINE-97122332; PubMed-8967952;
RA Huang B., Eberstadt M., Olejniczak E.T., Meadows R.P., Fesik S.W.;
RT "NMR structure and mutagenesis of the Fas (APO-1/CD95) death domain.";
RL Nature 384:638-641(1996).
CC -|- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
CC SUICIDE OF MATURE T-CELLS, OR BOTH.
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -|- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -|- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -|- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -|- DATABASE: NAME=PROW; NOTE=CD guide CD95 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd95.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M57454; AAA63174.1; -
DR EMBL; X63717; CAA45250.1; -
DR PIR; A40036; A40036.
DR PIR; S24543; S24543.
DR PDB; 1DDF; 12-NOV-97.
DR MIM; 134637; -
DR INTERPRO; IPR000488; -
DR INTERPRO; IPR001388; -
DR PFAM; PF00020; TNFR_c6; 2.
DR PFAM; PF00531; death; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00505; TNFR_NGFR_2; 2.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal;
KW 3D-structure.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 335 FASL RECEPTOR.
FT DOMAIN 17 173 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 174 190 POTENTIAL.
FT DOMAIN 191 335 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 47 166 3 X TNFR-CYS.
FT REPEAT 84 127 TNFR-CYS 1.
FT REPEAT 128 166 TNFR-CYS 2.
FT REPEAT 230 314 TNFR-CYS 3.
FT DOMAIN 118 118 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 335 AA; 37732 MW; 0139942535111410 CRC64;

Query Match 8.4%; Score 182; DB 1; Length 335;
Best Local Similarity 25.5%; Pred. No. 2.9e-06;
Matches 76; Conservative 34; Mismatches 114; Indels 74; Gaps 15;

QY 19 TTQETLPKPYHYDPETGHQLCDKCAPGYLKHQVTR-RKTLVCPDP-HSYTDSWHT 76
DB 42 TTVETQNLGELHHDGQFCHK----PCPPGERKARDCTVNGDEPDCVPCQSGKEYTKAHF 97
QY 77 SDCVYKSPVKELQSVKQEC---CNRTNHRVCECEEGRYLEIFELKHRSCPPGSGVWQAG 134
DB 98 SSKRCRRR-LCDEGHGLEVEINCTRTQNTKCRCKPNEFCNSTVC---EHCDP----- 145
```

```

Qy 135 TPERNTVCKGPDGFFSGET-SSKAPCIKHTNCSTFG---LLL-----IQKG 177
Db 146 -----CTKCEHGIIKECTLTSNTKCKEGRSRLGWLCLLLPIPLIVWVKREVK- 197
Qy 178 NATHDNVCSGNREATQKCGIDVTLCCEAFREFAVPTKIIPNWLSDLVDSLPCTK----- 231
Db 198 -----TCRKHKENQGSHPSTLNPE-----TVAINLSDVDLSKYITTTIAGVMTLSQVK 246
Qy 232 -----VNAESVERIKRRH--SSOEOTFOLLKWLKHONRDOEMVKKLIODI---DLC 277
Db 247 GFVRKNGVNEAKIDEIKNDNVQDTAEQVQLLRNWHQLHGRKREAYDTLIKDLKANLC 304

```

Search completed: March 1, 2001, 09:17:44
Job time: 262 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:42 ; Search time 79.26 Seconds
(without alignments)
343.530 Million cell updates/sec

Title: US-09-389-782A-2
Perfect score: 2179
Sequence: 1 MNKWLCCALLVLDIIEWTT.....OKLFLEMIGNOVQSVKISCL 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	314	14.4	459	2 I48854	gene murine tumour
2	312	14.3	461	1 A35356	tumor necrosis fac
3	312	14.3	474	2 B38634	tumor necrosis fac
4	270	12.4	289	2 A46515	B cell-associated
5	270	12.4	305	2 A46476	B cell-associated
6	261	12.0	277	2 A60771	B-cell activation
7	244	11.2	435	2 I54182	tumor necrosis fac
8	233	10.7	325	2 B43692	T2 protein - rabbi
9	220	10.1	326	1 GQVZML	T2 protein - myxom
10	217	10.0	349	2 D36858	gene G4R protein -
11	216	9.9	349	2 D72175	G2R protein - vari
12	215.5	9.9	348	2 T28623	hypothetical prote
13	199	9.1	427	1 GQHUN	nerve growth facto
14	197.5	9.1	425	1 A26431	nerve growth facto
15	184	8.4	416	1 JN0006	nerve growth facto
16	182	8.4	335	2 A40036	apoptosis-mediati
17	179.5	8.2	314	2 I37383	FAS soluble protei
18	177	8.1	256	2 B32393	T-cell antigen 4-1
19	167	7.7	255	2 I38426	4-1BB - human
20	166	7.6	255	2 JT0752	lymphocyte activat
21	165.5	7.6	461	2 JC4302	tumor necrosis fac
22	163	7.5	271	2 SI2783	OX40 antigen precu
23	161.5	7.4	454	1 GQMT1	tumor necrosis fac
24	161.5	7.4	454	2 I57826	tumor necrosis fac
25	160	7.3	455	1 GQHUT1	tumor necrosis fac
26	153	7.0	461	1 GQRUT1	tumor necrosis fac
27	152.5	7.0	272	2 I48700	gene ox40 protein
28	151	6.9	277	2 I37552	OX40 homolog - hum
29	147.5	6.8	324	2 JC2395	Fas antigen precu

30	143	6.6	327	2 A46484	apoptosis-mediati
31	142.5	6.5	2823	2 T23064	hypothetical prote
32	142.5	6.5	3102	2 T43291	laminin alpha chai
33	139	6.4	595	2 A42086	CD30 antigen precu
34	137.5	6.3	899	2 G02428	subtilisin-like pr
35	137.5	6.3	915	2 JC6148	subtilisin-like pr
36	133	6.1	1786	1 MMHUB1	laminin beta-1 cha
37	131.5	6.0	260	1 A46517	CD27 antigen precu
38	128.5	5.9	786	2 A48456	oocyst wall protei
39	128.5	5.9	1252	2 S36016	oocyst wall protei
40	128.5	5.9	1790	1 MMFEB1	laminin beta-1 cha
41	126.5	5.8	1548	2 S34583	serine proteinase
42	125	5.7	1680	2 A43434	furin (EC 3.4.21.7
43	122	5.6	1372	2 T25933	hypothetical prote
44	120.5	5.5	250	1 A49053	CD27 antigen precu
45	120.5	5.5	2918	2 A54105	fibrillin-2 precu

ALIGNMENTS

RESULT 1
I48854

gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C;Accession: I48854
R;Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.
Mamm. Genome 5, 726-727, 1994
A;Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.
A;Reference number: I48854; MUID:95178848
A;Accession: I48854
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-459 <RES>
A;Cross-references: EMBL:X76401; NID:9433830; PIDN:CAA53981.1; PID:9433831
C;Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
F;151-188/Domain: NGF receptor repeat homology <NGP>

Query Match 14.4%; Score 314; DB 2; Length 459;
Best Local Similarity 34.6%; Pred. No. 1.7e-14;
Matches 66; Conservative 20; Mismatches 79; Indels 26; Gaps 4;

QY	18	WTQETLPPKYL--HYDPETGH-----QLLCKCAPGYLKHCHVRRKTLQ	62
DB	2	WATGHTVPAQVVLTPYKPEPEGECIQSOEYDRKAQMCACPPGQYVKHFCNATSDTVCI	61
QY	63	VPCPDHSYTDSWHTSDCVYCSPVKELQSVKQECNTHNRVCECEGRYLEIEF-----	117
DB	62	ADCEASMYTQWNQFRCLSCSSSCSDQVETRACTKQQRNVCAACEAGRYCALKTHSGSC	121
QY	118	--CLKHRCPPGSGVQAGTPERNVTKCPDGFSSGETSSKAPCIKHTNGSTGLLIQ	175
DB	122	RQMRLSKCGPGFGVASSRPNPNVLCACAPGTFSDTSDTCDVRPHRIGS----ILAI	177
QY	176	KGNAHDNVCS	186
DB	178	PGNASTDVAVCA	188

RESULT 2
A35356

tumor necrosis factor receptor type 2 precursor - human
N;Alternate names: 75k tumor necrosis factor receptor
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A35356; A36475; A48416; A36007; A23666; B35010; I36094
R;Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dover, Science 248, 1019-1023, 1990
A;Title: A receptor for tumor necrosis factor defines an unusual family of cellular a
A;Reference number: A35356; MUID:90260639
A;Accession: A35356

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <SMI>
A:Cross-references: GB:M32315; NID:g189185; PIDN:AAA59929.1; PID:g189186
R:Kohn, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squires, Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally occurring
A:Reference number: A36475; MUID:91045991
A:Accession: A36475
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-195, 'R', 197-461 <KOH>
A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758
R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, M. Cytokine 2, 231-237, 1990
A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular, A:Reference number: A48416; MUID:91370690
A:Accession: A48416
A:Status: preliminary
A:Molecule type: protein
A:Residues: 23-461 <DEM>
A:Cross-references: GB:S63358; NID:g235648; PIDN:AAB19824.1; PID:g235649
A:Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBI:P:63371)
R:Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M. Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990
A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demonstration
A:Reference number: A36007; MUID:90349572
A:Accession: A36007
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 116-140, 'P', 142-195, 'R', 197-362, 'T', 364-461 <HEL>
A:Cross-references: GB:M35857; NID:g339751; PIDN:AAA63262.1; PID:g339752
R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus, M. J. Biol. Chem. 265, 20131-20138, 1990
A:Title: Purification and partial amino acid sequence analysis of two distinct tumor necrosis factor receptors
A:Reference number: A23666; MUID:91056048
A:Accession: A23666
A:Status: preliminary
A:Molecule type: protein
A:Residues: 23-40, 65-69, 136-141, 300-306 <LOE>
R:Engelmann, H.; Novick, D.; Wallach, D. J. Biol. Chem. 265, 1531-1536, 1990
A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Evidence
A:Reference number: A35010; MUID:90110215
A:Accession: B35010
A:Status: preliminary
A:Molecule type: protein
A:Residues: 27-31 <ENG>
R:Kuhnert, P.; Kemper, O.; Wallach, D. Gene 150, 381-386, 1994
A:Title: Cloning, sequencing and partial functional characterization of the 5' region of A:Reference number: I38094; MUID:95121934
A:Accession: I38094
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-37 <RES>
A:Cross-references: EMBL:X80021; NID:g666044; PIDN:CAA56324.1; PID:g825701
C:Genetics:
A:Gene: GDB:TNFR2
A:Cross-references: GDB:125914; OMIM:191191
A:Map position: lp36.2-1p36.2
A:Introns: 26/3
A:Note: the list of introns is incomplete
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-416/Product: tumor necrosis factor receptor type 2 #status experimental <MAT>
F:40-76/Domain: NGF receptor repeat homology <NG1>
F:78-119/Domain: NGF receptor repeat homology <NG2>
F:120-162/Domain: NGF receptor repeat homology <NG3>
F:164-201/Domain: NGF receptor repeat homology <NG4>
F:262-279/Domain: transmembrane #status predicted <TMN>
F:280-461/Domain: intracellular #status predicted <INT>

F:171,193/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 14.38; Score 312; DB 1; Length 461;
Best Local Similarity 34.78; Pred. No. 2.4e-14;
Matches 69; Conservative 19; Mismatches 85; Indels 26; Gaps 5;
Qy 8 ALLVLLDIIEWTQTETLPPK--YLHYDPTG-----QLLCDKCAPGTYLKOH 53
Db 9 ALAVGLEL--WAAHALPAQVAFPTPAPEPGSTCLREYYDQTQAMCCSKSPGGHAKVF 56
Qy 54 CTVRRKTLVPCPDHSYSDSHSTDSVYCVSPVCKELQSVKQECNRTNHRVCEEGRYL 113
Db 67 CTKTSDTVCDSCEDSTYTQLNNWVPECLSGSRSSDQVETQACTREQNRICTCRPGWYC 126
Qy 114 EI-----EFLCKHRSCPPGSGVQVAGTPERTVCKKCPDGGFFSGTSCAPCIKHTNCS 167
Db 127 ALSQKQECRLCAPLRKCRPGFVARPGTETSDVVKCPACPGTFSNTTSDICRPHQICN 186
Qy 168 TFGLLLIQKGNATHDNVCS 186
Db 187 VVAI-----PGNASMDAVCT 201
RESULT 3
B38634
tumor necrosis factor receptor type 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: B38634; A40254; S54816
R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen, Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991
A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis factor A:Reference number: A38634; MUID:91187885
A:Accession: B38634
A:Molecule type: mRNA
A:Residues: 1-474 <LEW>
A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; J Mol. Cell. Biol. 11, 3020-3026, 1991
A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors f A:Reference number: A40254; MUID:91246168
A:Accession: A40254
A:Molecule type: mRNA
A:Residues: 1-474 <GOO>
A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R:Kisssonerghis, M.; Fallowes, R.; Feldmann, M.; Chernajovsky, Y. submitted to the EMBL Data Library, May 1995
A:Description: Characterization of the promoter region of the murine p75-TNF receptor A:Reference number: S54816
A:Accession: S54816
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <KIS>
A:Cross-references: EMBL:X87128; NID:g809043; PIDN:CAA06018.1; PID:g809044
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C:Keywords: cytokine receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-474/Product: tumor necrosis factor receptor type 2 #status predicted <MAT>
F:40-77/Domain: NGF receptor repeat homology <NG1>
F:79-120/Domain: NGF receptor repeat homology <NG2>
F:166-203/Domain: NGF receptor repeat homology <NG4>

Query Match 14.38; Score 312; DB 2; Length 474;
Best Local Similarity 34.38; Pred. No. 2.4e-14;
Matches 70; Conservative 22; Mismatches 84; Indels 28; Gaps 5;
Qy 5 LCCALLVLLDIIEWTQTETLPPKYL--HYDPTG-----QLLCDKCAPGTY 49
Db 6 LWVALVFELQL--WATGHTVPAQVVLTPYKPEPGVECOISQVEYDRKAQMCCKAPPGQY 63
Qy 50 LKQHCTVRRKTLVPCPDHSDTSDSWHTSDEVCYCSVPCKELQSVKQECNRTNHRVCECEE 109

```
Db 64 VKHFNKTSVTCADCEASMYTQWNOFRCLSCSSCTTDOVEIRACTRQONRVACAEA 123
QY 110 GRYLEIEF-----CLKHRSPPGSGVQVAGTPERTNVCKCPDGFSGFSTSSKAPCIK 162
Db 124 GRVCAKTHSGSRQCRMLKSGPGFGVASSRAPNGVNLCKACAPGFTSDTSSDVCRP 183
QY 163 HTNCSTFGLLLIQGNATHDNVCS 186
Db 184 HRICS-----ILAIPGNASTDAVCA 203

RESULT 4
A64515
B cell-associated surface molecule CD40, short splice form - mouse
C:Species: Mus musculus (house mouse)
C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 26-May-2000
C:Accession: A46515
R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockayne,
J. Immunol. 149, 3921-3926, 1992
A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.
A:Reference number: A46515; MUID:93094586
A:Accession: A46515
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-289 <GRI>
A:Cross-references: GB:M83312; NID:g1553058; PIDN:AAB08705.1; PID:g1553059; GB:M94126; N
A:Experimental source: BALB/c, liver
A:Note: sequence extracted from NCBI backbone (NCBIP:120357)
C:Comment: For an alternative splice form, see PIR:A46476
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: alternative splicing
F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 12.4%; Score 270; DB 2; Length 289;
Best Local Similarity 34.5%; Pred. No. 1.2e-11;
Matches 57; Conservative 26; Mismatches 68; Indels 14; Gaps 5;
QY 27 KYLHYDPETGHQLLCKCAPGTYLKQCHTNRVRRKTLVCPDHSYDTSWHTSDEC---VYC 83
Db 30 QYLH-DGQ-----CCDLCPGSRULTSHTALEKTQCHPCDGSFGESFAQWNRREIRCHQHRHC 83
QY 84 SPVCKELQSVKQECNRTHNRVCEEGRYL---EIEFLKHRSPGSGVQVAGTPERTN 140
Db 84 EP--NOGLRVKKEGTAEISDVCTCKEGQCHTSKDCEACAHQTPCIPGFGVEMATETD 141
QY 141 VCKKCPDGFSGFSTSSKAPCIKHTNCSTFGLLLIQGNATHDNVNC 185
Db 142 VCHPCPVGFFSNQSSLFKCYPTWTSCEDKNLEVLQKGTSTQTNVIC 186

RESULT 5
A46476
B cell-associated surface molecule CD40, long splice form - mouse
C:Species: Mus musculus (house mouse)
C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 26-May-2000
C:Accession: A46476
R:Torres, R.M.; Clark, E.A.
J. Immunol. 148, 620-626, 1992
A:Title: Differential increase of an alternatively polyadenylated mRNA species of murine
A:Reference number: A46476; MUID:92105763
A:Accession: A46476
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-305 <TOR>
A:Cross-references: GB:M83312; NID:g1553058
A:Note: sequence extracted from NCBI backbone (NCBIN:75206, NCBIP:75207)
A:Note: this translation is not annotated in GenBank entry MUSCD40A, release 113.0
C:Comment: For an alternative splice form, see PIR:A46515
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: alternative splicing; transmembrane protein
F:105-144/Domain: NGF receptor repeat homology <NGF>
```

```
Query Match 12.4%; Score 270; DB 2; Length 305;
Best Local Similarity 34.5%; Pred. No. 1.2e-11;
Matches 57; Conservative 26; Mismatches 68; Indels 14; Gaps 5;
QY 27 KYLHYDPETGHQLLCKCAPGTYLKQCHTNRVRRKTLVCPDHSYDTSWHTSDEC---VYC 83
Db 30 QYLH-DGQ-----CCDLCPGSRULTSHTALEKTQCHPCDGSFGESFAQWNRREIRCHQHRHC 83
QY 84 SPVCKELQSVKQECNRTHNRVCEEGRYL---EIEFLKHRSPGSGVQVAGTPERTN 140
Db 84 EP--NOGLRVKKEGTAEISDVCTCKEGQCHTSKDCEACAHQTPCIPGFGVEMATETD 141
QY 141 VCKKCPDGFSGFSTSSKAPCIKHTNCSTFGLLLIQGNATHDNVNC 185
Db 142 VCHPCPVGFFSNQSSLFKCYPTWTSCEDKNLEVLQKGTSTQTNVIC 186

RESULT 6
A60771
B-cell activation protein CD40 precursor - human
N:Alternate names: B-cell surface antigen Bp50
C:Species: Homo sapiens (man)
C>Date: 03-Jun-1993 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: S04460; A60771
R:Stamenkovic, I.; Clark, E.A.; Seed, B.
EMBO J. 8, 1403-1410, 1989
A:Title: A B-lymphocyte activation molecule related to the nerve growth factor receptor
A:Reference number: S04460; MUID:89356608
A:Accession: S04460
A:Molecule type: mRNA
A:Residues: 1-277 <STA>
A:Cross-references: EMBL:X60592; NID:g29850; PIDN:CAA43045.1; PID:g29851
R:Braesch-Andersen, S.; Paulie, S.; Kono, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.
J. Immunol. 142, 562-567, 1989
A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-
A:Reference number: A60771; MUID:89093941
A:Accession: A60771
A:Molecule type: protein
A:Residues: 21-50 <BRA>
A:Experimental source: Burkitt lymphoma cell line Raji
C:Genetics:
A:Gene: GDB:CD40
A:Cross-references: GDB:215268; OMIM:109535
A:Map position: 20q12-20q13.2
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane prot
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-277/Product: B-cell activation protein CD40 #status experimental <MAT>
F:21-193/Domain: extracellular #status predicted <EXT>
F:194-215/Domain: transmembrane #status predicted <TM>
F:216-277/Domain: intracellular #status predicted <CYT>
F:153,180/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 12.0%; Score 261; DB 2; Length 277;
Best Local Similarity 33.9%; Pred. No. 4.6e-11;
Matches 58; Conservative 23; Mismatches 74; Indels 16; Gaps 4;
QY 29 LHYDPETG-----HOLLCDKCAPGTYLKQCHTNRVRRKTLVCPDHSYDTSWHTSDEC 80
Db 18 VHEPPTACREKQYVLINSQCCLQCPGKLVSDCTETFTETCLCGESEFLDTWNRTHC 77
QY 81 ---VYCSPVCKELQSVKQECNRTHNRVCEEGRYL---EIEFLKHRSPGSGVQVAG 134
Db 78 HQHYKDFPNLGL--LRVQKGTSETDICTCEBWHCTSEACESCVLHRSRSPGFGVKQIA 135
QY 135 TPERNTVCKKCPDGFSGFSTSSKAPCIKHTNCSTFGLLLIQGNATHDNVNC 185
Db 136 TGVSDTICEPCPVGFFSNVSAFEKCHPWTSCETNDLVVQVQAGNKTDVVC 186
```

RESULT 7
154182
tumor necrosis factor receptor 2-related protein - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Mar-2000
C:Accession: 154182
R:Baens, M.; Chaffanet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
Genomics 16, 214-218, 1993
A:Title: Construction and evaluation of a hncDNA library of human 12p transcribed sequences
A:Reference number: 154182; MUID:93252381
A:Accession: 154182
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-435 <RES>
A:Cross-references: GB:L04270; NID:g339761; PIDN:AAA36757.1; PID:g339762
C:Genetics:
A:Gene: GDB:LTBR
A:Cross-references: GDB:1230195; OMIM:600979
A:Map position: 12p13.3-12p13.1
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match 11.2%; Score 244; DB 2; Length 435;
Best Local Similarity 26.9%; Pred. No. 1.1e-09;
Matches 54; Conservative 32; Mismatches 93; Indels 22; Gaps 6;

QY 9 LVLVLDIETWTTQETLPP-----KYLHYDPETGHQLLCKCAPGTVLKQKCTVR 57
DB 18 VLGLFGLLAASQPAVPYASENQTCROEKEYEYEPQ--HRICCSRCPPGIYVSAKCSRI 75
QY 58 RKLTCVPCPDHSDYTSWHTSDCYVSPCKELQSVKQ--ECNRTHNRVCEEGRY--- 112
DB 76 RDTVCATCAENSYNEHNYLTICQLCRP--CDPVMGLEETAPCTSKRTKTCRCQCPGMFCAA 134
QY 113 --LETEFLKLRSCPPGS--GVVQAGTPERTNVCKKCPDGFSGETSSKAPCTKHNCSTF 169
DB 135 WALECTHCELLSDCPPTGAELKDEVGNNHCVPCKAGHFONTSSPACQPHTRCENQ 194
QY 170 GLLLIQKGNATHDNVCSGNRE 190
DB 195 GLVEAAPGTAQSDTTCKNPLE 215

RESULT 8
B43692
T2 protein - rabbit fibroma virus
C:Species: rabbit fibroma virus, Shope fibroma virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: B43692
R:Upton, C.; DeLange, A.M.; McFadden, G.
Virology 160, 20-30, 1987
A:Title: Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric
A:Reference number: B43692; MUID:87321103
A:Accession: B43692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <UPT>
A:Cross-references: GB:M17433
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>

Query Match 10.7%; Score 233.5; DB 2; Length 325;
Best Local Similarity 31.7%; Pred. No. 4.4e-09;
Matches 59; Conservative 19; Mismatches 91; Indels 17; Gaps 5;

QY 36 GHQ-----LLCDKCAPGYLKQKCTVRRTKLCVPCPDHSDYTSWHTSDCYVSPCKEL 90
DB 30 GHDEKDGKGLCCASCHPGFYASRLCGPGSNTVCSPCEDGTFTASTNHAPACVSCRCGPCTGH 89
QY 91 QSVKQECNRTHNRVCEEGRYLEIE-----FCLKHRSCPPSGVVGAGTPERTNVCKK 144

DB 90 LSEQPCDRTHRVNCSTGNVCLLKQNGCRICAPQTKCPAGYG--VSGHTRAGTFLCEK 148
QY 145 CPDGFSGETSKAPCIKHTNCSTFGLLLIQKGNATHNVCSGNREATOKCGIDVTL--- 201
DB 149 CPPHTYSLSLSPTEPCGTSFNKISVGFNLVYPV-NETSCSTTTAGHNEVIKTEFTVLTNYT 207
QY 202 -CEEAF 206
DB 208 DCDPVF 213

RESULT 9
GQVZML
T2 protein - myxoma virus (strain Lausanne)
C:Species: myxoma virus
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 18-Jun-1999
C:Accession: A40566
R:Upton, C.; Macen, J.L.; Schreiber, M.; McFadden, G.
Virology 184, 370-382, 1991
A:Title: Myxoma virus expresses a secreted protein with homology to the tumor necrosis
A:Reference number: A40566; MUID:91335768
A:Accession: A40566
A:Molecule type: DNA
A:Residues: 1-326 <UPT>
A:Cross-references: GB:M95181; GB:M37976; NID:g332309; PIDN:AAA46632.1; PID:g332310
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
C:Keywords: glycoprotein
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>
F:66,181,205,238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.1%; Score 220; DB 1; Length 326;
Best Local Similarity 29.1%; Pred. No. 3.8e-08;
Matches 53; Conservative 24; Mismatches 85; Indels 20; Gaps 5;

QY 39 LLCDKCAPGYLKQKCTVRRTKLCVPCPDHSDYTSWHTSDCYVSPCKELQVKBECN 98
DB 38 LCCTSCPPGSYASRLCGPGSDTVCSPCNKETFTASTNHAPACVSCRCRGTGHLSESQSCD 97
QY 99 RTHNRVCEEGRYLEIE-----FCLKHRSCPPSGVVGAGTPERTNVCKKCPDGFSG 152
DB 98 KTRDRVCDCSAGNYCLLKQEGSCRICAPKTKCPAGYG--VSGHTRTGDVLTCKPRTYSD 156
QY 153 ETSSKAPCIKHTNCSTFGLLLIQK-----NATHDNVCSGNREATOKCGIDVTL---CEE 204
DB 157 AVSSTETC-----TSSFNKISVEFNLYPVNDTSCSTTTAGPNEVVKTSFSTLHTDCDP 211
QY 205 AF 206
DB 212 VF 213

RESULT 10
D36858
gene G4R protein - variola virus
N:Alternate names: B28R protein (COP)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
C:Accession: D36858; S46888; S32385; S35987
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Description: not shown.
A:Reference number: A36859
A:Accession: D36858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <BLI>
A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49137.1; PID:g457087
A:Experimental source: strain India-1967, ssp. major, isolate Ind3
R:Kolykhalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F
submitted to the EMBL Data Library, April 1992
A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H

C:Genetics:
A:Gene: G2R
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

Query Match 9.9%; Score 216; DB 2; Length 349;
Best Local Similarity 23.2%; Pred. No. 7.7e-08;
Matches 92; Conservative 51; Mismatches 169; Indels 84; Gaps

QY 4 WLCCALLVLIDIEWTQTETLPKYLHYDPE-TGHOLLCDKCAPGYLKQHCTVRRKTLIC 175
DB 11 FLSCIINGRDAAPYT-----PENGCKKDEYKRHNLCCLSCPPGTIYASRLCDSKNTCT 65

QY 63 VPCPDHSYTDSWHTSDVCYSPVKELQSVKQECNRTHNRVCECEBGRYLEI-----FE 116
DB 66 TPCGSGTFTSRNNHLPACLSCNGRCNSNOVETRSCNTTHNRICECPGYCYLLKGSSGCK 125

QY 117 FCLKHSRCPGSGVQVQAGTPERTNVCKKCPDGFSETSKAPCIKHTNCSIFGLLIQ- 175
DB 126 ACVSQTKCGIGYG-VSGHTSVGDVICSPCGFGTYSYTVSDTKECPVN-NTFNYIDVEI 183

QY 176 KGNATHDNVCSGNREATQKGDIVTLCEEAFFRAVPTKIIPNLSVLVDSLPGTKVNAE 235
DB 184 TLPVNDTSC-----RTTTGLSE-----SILTSEL----- 210

QY 236 SVERIKRRHSS-----QEQTQLLK-----LWKHQNROQEMVKKIQIDIDLCSSVQH 284
DB 211 ---TITMHTDCNPVFEFVSVLNKVATSGPFTGENRYQNISK-----VCTLNFELK 260

QY 285 LGHSNLTTEOLLALMSLPCKKISPEIERTRTKCKSSQOLLKLLSLWRKIKNGDQDTLAG 344
DB 261 CNKGSFCKQLTKAKND--DGMMSHSETVTLAGDCLSSVDIVILYSNTNAQDYETDIS- 317

QY 345 LMVALKHL--KTSHEP-----KTVTHSLRKTMREL 372
DB 318 --YRVGNVLDDSHMPGSCDIHKLITNS--KPTRFL 349

RESULT 12
T28623
Hypothetical protein G2R - variola major virus
C:Species: variola major virus
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T28623
R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight
Nature 366, 748-751, 1993
A:Title: Potential virulence determinants in terminal regions of variola
A:Reference number: Z20488; MUID:94088747
A:Accession: T28623
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-348 <N>
A:Cross-references: EMBL:L22579; NID:G623595; PIDN:AAA60933.1; PID:939102
A:Experimental source: strain Bangladesh 1975
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

Query Match 9.9%; Score 215.5; DB 2; Length 348;
Best Local Similarity 23.1%; Pred. No. 8.4e-08;
Matches 87; Conservative 47; Mismatches 156; Indels 87; Gaps

QY 31 YDPETG-----HOLLCDKCAPGYLKQHCTVRRKTLCPDHSYTDSWHTSDECV 81
DB 24 YTPNGKCKDEYKRHNLCCLSCPPGTIYASRLCDSKNTCTCTCGSGTFTSRNNHLPACL 83

QY 82 YCSPVKELQSVKQECNRTHNRVCECEBGRYLEI-----EFCIKHSRCPGSGVQVQAGT 135
DB 84 SCNGRCNSNOVETRSCNTTHNRICECPGYCYLLKGSSGCKACVSQTKCGIGYG-VSGHT 142

QY 136 PERTNVCKKCPDGFSETSKAPCIKHTNCSIFGLLIQ-KGNATHDNVCSGNREATQK 194
DB 143 SVGDVICSPCGFGTYSHTVSSAKCEPVPN-NTFNYIDVEITILYPVNDTSC-----RTTT 197

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:20:05 ; Search time 135.68 seconds
(without alignments)
345.542 Million cell updates/sec

Title: US-09-389-782a-5
Perfect score: 2240
Sequence: 1 ETFFPKYLHYDEETSHQLLC.....VMHEALHNHYTKSLSPG 400

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1013.5	45.2	401	4 000300	000300 homo sapien
2	1013	45.2	372	4 09UHP4	09UHP4 homo sapien
3	893.5	39.9	401	11 008712	008712 mus musculus
4	877	39.2	401	11 008727	008727 rattus norv
5	827	36.9	437	11 09RIA4	09RIA4 mus musculus
6	433	19.3	300	4 095407	095407 homo sapien
7	405.5	18.1	302	13 09PUS0	09pus0 salvelinus
8	345	15.4	459	11 062327	062327 mus musculus
9	343.5	15.3	439	4 016042	016042 homo sapien
10	338	15.1	482	11 088734	088734 mus musculus
11	297	13.3	655	4 075509	075509 homo sapien
12	280.5	12.5	684	13 090544	090544 ginglymosto
13	277.5	12.4	384	4 09UP60	09up60 homo sapien
14	273.5	12.2	416	4 09NPP6	09npp6 homo sapien
15	258.5	11.5	616	4 09Y606	09y606 homo sapien
16	256.5	11.5	625	11 035305	035305 mus musculus
17	241	10.8	349	12 057099	057099 monkeypox v
18	239	10.7	349	12 057291	057291 monkeypox v
19	239	10.7	349	12 057100	057100 monkeypox v

20	239	10.7	349	12	057101	057101 monkeypox v
21	239	10.7	349	12	057102	057102 monkeypox v
22	235.5	10.5	348	12	057277	057277 monkeypox v
23	235.5	10.5	348	12	057103	057103 monkeypox v
24	235.5	10.5	348	12	057108	057108 monkeypox v
25	229	10.2	348	12	057112	057112 variola vir
26	229	10.2	348	12	085407	085407 variola vir
27	226.5	10.1	349	12	057284	057284 camelpox vi
28	226.5	10.1	349	12	057098	057098 camelpox vi
29	226	10.1	349	12	057110	057110 variola vir
30	226	10.1	349	12	057111	057111 variola vir
31	226	10.1	349	12	089118	089118 variola vir
32	226	10.1	349	12	089098	089098 variola vir
33	223	10.0	350	12	057116	057116 cowpox viru
34	222.5	9.9	349	12	057097	057097 camelpox vi
35	222.5	9.9	350	12	057123	057123 cowpox viru
36	222.5	9.9	355	12	085308	085308 cowpox viru
37	221	9.9	349	12	057109	057109 variola vir
38	220.5	9.8	349	12	057305	057305 cowpox viru
39	219	9.8	326	12	057122	057122 cowpox viru
40	218.5	9.8	360	12	057118	057118 cowpox viru
41	217.5	9.7	326	12	057120	057120 cowpox viru
42	217.5	9.7	351	12	073559	073559 cowpox viru
43	216.5	9.7	351	12	057117	057117 cowpox viru
44	212.5	9.5	347	12	057115	057115 cowpox viru
45	210.5	9.4	316	12	057092	057092 ectromelia

ALIGNMENTS

RESULT 1
000300
ID AC 000300 PRELIMINARY; PRT; 401 AA.
AC 000300; 060236;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11b).
GN TNFRSF11B OR OPG OR OCIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RA *Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density.";
RL Cell 89:309-319(1997).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG FIBROBLAST;
RX MEDLINE=98151033; PubMed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
RA Tsuda E., Morinaga T., Higashio K.;
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RL osteoclastogenesis in vitro.";
RL Endocrinology 139:1329-1337(1998).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; PubMed=9688283;

```

RA Moringa T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human
RT osteoprotegerin/osteoclastogenesis-inhibitory factor.";
RL Eur. J. Biochem. 254:685-691(1998).
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; AB002146; BAA25910.1; -.
DR EMBL; AB008822; BAA32076.1; -.
DR EMBL; AB008821; BAA32076.1; JOINED.
DR EMBL; U94332; AAB53709.1; -.
DR HSSP; P25942; 1CDF.
DR MIM; 602643; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
DR PRODOM; PD00071; -.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21
FT CHAIN 22 401 By SIMILARITY.
FT DOMAIN 23 183 OSTEOPROTEGERIN.
FT REPEAT 23 63 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 45.2%; Score 1013.5; DB 4; Length 401;
Best Local Similarity 68.7%; Pred. No. 3.1e-81;
Matches 193; Conservative 12; Mismatches 37; Indels 39; Gaps 5;

Qy 1 ETFPPKYLHYDEETSHQLLCKDKCPGTYLKQHCTAKKTVKVCAPCPDHYTDSWHTSDECL 60
Db 22 ETFPPKYLHYDEETSHQLLCKDKCPGTYLKQHCTAKKTVKVCAPCPDHYTDSWHTSDECL 81
Qy 61 YCSPVKELQYVQKQECNTRNHRVCECKEGRYLEIEFCLKLRHSCPPGFGVQAGTPERNTV 120
Db 82 YCSPVKELQYVQKQECNTRNHRVCECKEGRYLEIEFCLKLRHSCPPGFGVQAGTPERNTV 141
Qy 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTQKVDKTHTC 180
Db 142 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTQK 194
Qy 181 PCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTVVVDVSHEDPEVKFNWYVDGVEVHN 240
Db 195 --CGIDVTLCCEAFAFRFAVPTK-----FTPNWLSVLVD-----NLPCKVNA 234

```

```

Qy 241 AKTKPREQYNSTYRVSVLTVLHODWLNKYEYKCKVSNKA 281
Db 235 ESVERIKRQHSSEQOTFOLLKL-----W-----KHQNK 263

RESULT 2
Q9UHP4 PRELIMINARY; PRT; 372 AA.
AC Q9UHP4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE OSTEOPROTEGERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).
DR EMBL; AF134187; AAF20168.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
FT NON_TER 1
FT SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 45.2%; Score 1013; DB 4; Length 372;
Best Local Similarity 73.4%; Pred. No. 3.2e-81;
Matches 190; Conservative 10; Mismatches 35; Indels 24; Gaps 5;

Qy 1 ETFPPKYLHYDEETSHQLLCKDKCPGTYLKQHCTAKKTVKVCAPCPDHYTDSWHTSDECL 60
Db 1 ETFPPKYLHYDEETSHQLLCKDKCPGTYLKQHCTAKKTVKVCAPCPDHYTDSWHTSDECL 60
Qy 61 YCSPVKELQYVQKQECNTRNHRVCECKEGRYLEIEFCLKLRHSCPPGFGVQAGTPERNTV 120
Db 61 YCSPVKELQYVQKQECNTRNHRVCECKEGRYLEIEFCLKLRHSCPPGFGVQAGTPERNTV 120
Qy 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTQKVDKTHTC 180
Db 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTQK 173
Qy 181 PCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTVVVDVSHEDPEVKFNWYVDGVEVHN 240
Db 174 --CGIDVTLCCEAFAFRFAVPTK-----FTPNWLSVLVD---AESVE--- 217
Qy 241 AKTKPREQYNSTYRVSV 259
Db 218 -RIKROHSSEQOTFOLLKL 235

RESULT 3
Q08712 PRELIMINARY; PRT; 401 AA.
AC Q08712; 070202;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF).
GN TNFRSF11B OR OPB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-KIDNEY;
 RX MEDLINE=97262071; PubMed=9108485;
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
 RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
 RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
 RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
 RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
 RA Suggs S., Boyle W.J.,
 RA "Osteoprotegerin: a novel secreted protein involved in the regulation
 of bone density.";
 RT Cell 89:309-319(1997).
 RL [2]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/OLA, AND NIH SWISS;
 RX MEDLINE=98382527; PubMed=9714833;
 RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,
 RA Higashio K.;
 RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
 gene and its expression in embryogenesis.";
 RL Gene 215:339-343(1998).
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
 OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
 SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
 OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
 STROMAL CELLS AND OSTEOCLAST PROGENITORS.
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,
 BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND
 PLACENTA. NOT DETECTED IN SPLEEN.
 CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT
 DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY
 15 TO DAY 17.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL; U94331; AAB53708.1; -;
 DR EMBL; AB013898; BAA28269.1; -;
 DR EMBL; AB013903; BAA33388.1; -;
 DR EMBL; AB013899; BAA33388.1; JOINED.
 DR EMBL; AB013900; BAA33388.1; JOINED.
 DR EMBL; AB013901; BAA33388.1; JOINED.
 DR EMBL; AB013902; BAA33388.1; JOINED.
 DR HSSP; P25942; 1CDF.
 DR MGD; MGI:109587; Opg.
 DR INTERPRO: IPR000488; -;
 DR INTERPRO: IPR001368; -;
 DR PFAM; PF00020; TNFR_c6; 3.
 DR PROSITE; P500652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE; P50017; DEATH_DOMAIN; 1.
 DR PROSITE; P50050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -; 1.
 KW Glycoprotein; Repeat; Cytokine; Signal.
 FT SIGNAL
 FT CHAIN 1 21
 FT CHAIN 22 401
 FT DOMAIN 23 201
 FT REPEAT 23 63
 FT REPEAT 24 63
 FT REPEAT 64 106
 FT REPEAT 107 143
 FT REPEAT 144 201
 FT REPEAT 201 365
 FT DOMAIN 306 365
 FT DISULFID 41 54
 FT DISULFID 44 62
 FT DISULFID 65 80
 FT DISULFID 83 97
 FT DISULFID 87 105
 FT DISULFID 118 142
 FT DISULFID 145 160
 FT CARBOHYD 98 98
 FT CARBOHYD 165 165
 FT CARBOHYD 178 178
 FT CARBOHYD 289 289
 FT VARIANT 138 138
 FT R -> P (IN STRAINS 129/OLA AND NIH
 FT SWISS).

FT VARIANT 161 161 I -> R (IN STRAINS 129/OLA AND NIH
 FT SWISS).
 FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH
 FT SWISS).
 FT VARIANT 288 288 S -> A (IN STRAINS 129/OLA AND NIH
 FT SWISS).
 FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH
 FT SWISS).
 SQ SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;
 Query Match 39.9%; Score 893.5; DB 11; Length 401;
 Best Local Similarity 47.8%; Pred. No. 1.1e-70;
 Matches 191; Conservative 35; Mismatches 107; Indels 67; Gaps 9;
 QY 1 ETFPKYLHDEETSHQLCDKCPGPGTYLKQHCTAKWTKVCAPCPDHVYTDSWHTSDECL 60
 DB 22 ETLPKYLHVDPTGHQLCDKCAPGTYLKQHCTVRRKTLCPDPHSYTDSWHTSDECY 81
 QY 61 YCSPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHSRCPGPGVQAGTPERTV 120
 DB 82 YCSPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHSRCPGPGVQAGTPERTV 141
 QY 121 CKRCPDGFNSSTSSKAPCRKHTNCVFGLLLTQKGNATHDNCISGSESTQKVDKTHTC 180
 DB 142 CKKCPDGFNSSTSSKAPCRKHTNCVFGLLLTQKGNATHDNCISGSESTQKVDKTHTC 194
 QY 181 PCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEHN 240
 DB 195 --CGIDVTLCCEAFFFAVPTK-----IIPNMLSVLVD---SLPGTKVN--AESVE-- 238
 QY 241 AKTKPREQNSYRVVSVLTVLHQLDNLNGKEY-----CKKVS----- 278
 DB 239 -RKRRHSSQEQFQLLKWKHQNRQEMVKIIQDIDLCESSVQRHLGHSNLTTEQLIA 297
 QY 279 -----NKALPAPIEKTISKAKGPQPVYTL-----PPSRDELTKNQVSLTCLVKGFT 327
 DB 298 LMESLPKGIKISPEIERTKTRCTCKSSQELLKLLSLRWKNGDQDTLKLGLMVALKHLKSHF 357
 QY 328 PSDIANEWESNGQPNENYKTPPVLDSDGFFLYSKLTVD 367
 DB 358 PKTVT-----HSLRKTMRFLSHFTMYRLYQKLFLE 387
 RESULT 4
 R08727 PRELIMINARY; PRT; 401 AA.
 ID 008727 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
 DE (OCIF).
 GN TNFRSF11B OR OPG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OC NCBI_taxid=10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=INTESTINE;
 RX MEDLINE=97262071; PubMed=9108485;
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
 RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
 RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
 RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
 RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
 RA Suggs S., Boyle W.J.,
 RT "Osteoprotegerin: a novel secreted protein involved in the regulation
 of bone density.";
 RL Cell 89:309-319(1997).
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
 OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY

```
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94330; AAB53707.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21
FT CHAIN 22 401 BY SIMILARITY.
FT DOMAIN 23 201 OSTEOPROTEGERIN.
FT REPEAT 23 63 4 X TNFR-CYS.
FT REPEAT 64 106 TNFR-CYS 1.
FT REPEAT 107 143 TNFR-CYS 2.
FT REPEAT 144 201 TNFR-CYS 3.
FT DOMAIN 306 365 TNFR-CYS 4.
FT DISULFID 41 54 DEATH DOMAIN.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
SQ SEQUENCE 401 AA; 46192 MW; FEC6A31F1D4E573A CRC64;

Query Match 39.2%; Score 877; DB 11; Length 401;
Best Local Similarity 63.7%; Pred. No. 3.2e-69;
Matches 165; Conservative 18; Mismatches 52; Indels 24; Gaps 5;

Qy 1 ETFPPKYLHYDEESHQILLCKPCPTGTYLKQHCYAKWTVCACPDHYVTSWHTSDCL 60
Db 22 ETFPPKYLHYDPETGRQLLCKPCAGTYLTKQHCYAKWTVCACPDHYVTSWHTSDCV 81

Qy 61 YCSPVKELQVQKBCNTHNRVCEGRYLETEFCLKHSCPPGFGVQAGTPERNIV 120
Db 82 YCSPVKELQVQKBCNTHNRVCEGRYLETEFCLKHSCPPGFGVQAGTPERNIV 141

Qy 121 CKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTQKVDKTHTC 180
Db 142 CKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTQKVDKTHTC 194

Qy 181 PPCPAPELLGSPVLEFPKPKDILMISRTPEVTCVVDVSHEDPEVKNFYVDGVEVHN 240
Db 195 --GIDVTLCGEAFRAVPTK-----IIPNLSVLVD---SLPGTKVN---AESVE--- 238

Qy 241 AKTKPREQYNSTYRVVSV 259
Db 239 -RIKRRHSSQEQTFQLKL 256

RESULT 5
Q9RIA4 PRELIMINARY; PRT; 437 AA.
AC Q9RIA4:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -.
DR HSSP; P01842; 7FAB.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; Ig; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 36.9%; Score 827; DB 11; Length 437;
Best Local Similarity 60.9%; Pred. No. 8.8e-65;
Matches 145; Conservative 44; Mismatches 39; Indels 10; Gaps 3;

Qy 170 STQKVDKTHT-----CPP--CPAPELLGSPVFLPPPKDPTLMISRTPEVTCVVDVSH 222
Db 202 SSTKVDKKIVPRDCGCKPCICTVPEV---SSVFIEPPKPKDVLITLTPKVTCTVVDISK 258

Qy 223 EDPEVKNWYDGVVEVHNKTPREEQYNSTYRVVSVLTVLHODWLNKGEYCKYKSNKAL 282
Db 259 DPEVQFSWFDVDDVEHTAQTPREEQNFSTRSELPIMHQDWLNKGEYCKYKSNKAL 318

Qy 283 PAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOPE 342
Db 319 PAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOPE 378

Qy 343 NNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNFSCSVMHEALHNHYTQKSLSLSPG 400
Db 379 ENYKNTQPIMDTGSFYFYSKLVQKSNWEAGNTFTCSVLHEGLHNHTKNSLSPG 436

RESULT 6
O95407 PRELIMINARY; PRT; 300 AA.
AC O95407:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN DCR3 OR TR6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99087326; PubMed=9872321;
RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel P.C.,
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA Goddard A.D., Botstein D., Ashkenazi A.;
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
RT colon cancer.";
RL Nature 396:699-703(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=99253915; PubMed=10318773;
RA Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT "A newly identified member of tumor necrosis factor receptor
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL J Biol. Chem. 274:13733-13736(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
```



```

RX MEDLINE-20122600; PubMed-10655513;
RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA Soderman A., Galloway S.M., Liu O., Austin C.P., Caskey C.T.;
RT "Overexpression of M88/DCR3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RT cluster."
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR EMBL; AF104419; AAD03056.1; -.
DR EMBL; AF134240; AAD29688.1; -.
DR EMBL; AF217796; AAF35244.1; -.
DR EMBL; AF217793; AAF33685.1; -.
DR EMBL; AF217794; AAF33686.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_NGFR_1; UNKNOWN_1..
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1..
DR PROSITE; PS00186; EGF_2; UNKNOWN_1..
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match
Best Local Similarity 19.3%; Score 433; DB 4; Length 300;
Matches 69; Conservative 29; Mismatches 63; Indels 0; Gaps 0;

QY 5 PKYLHDETSHQLLCDKCPGGTYLKQHTAKWKTVCAPCPDHYHSDSWHTSDECLYCSP 64
DB 34 PTYFWRDAETGELVLAQCQPPGTFVORPCRRDSPTTCGCPGPRHYTFQWYLERCRYCNV 93
QY 65 VKLELYVQECNRTNHRVCECKEGRYL--ETFECLKHRSCPPGFGVQAGTPERTNTVCKRC 124
DB 94 LCGEREERACHATNHRACRTGTFAGHAGFCLEHASCPGAGVIAPGTPSQNTQCQC 153
QY 125 PDGFFSNETSAPCRKHTNCSVFGLLLTKQGNATHDNICS 165
DB 154 PPGTFSSASSSSQCPHRCNTALGLALNVPYSGSSHDTLCT 194

RESULT 7
ID Q9PUS0 PRELIMINARY; PRT; 302 AA.
AC Q9PUS0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE DECOY RECEPTOR.
OS Salvelinus fontinalis (Brook trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN 1;
RP SEQUENCE FROM N.A.
RA Robe J., Goetz F.W.;
RT "A tumor necrosis factor receptor homolog is up-regulated in the brook
RT trout (Salvelinus fontinalis) ovary at the completion of ovulation.";
RL Biol. Reprod. 0:0-0(1999).
DR EMBL; AF156738; AAD56428.1; -.
DR HSSP; P19438; 1EXT.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_NGFR_1; UNKNOWN_1..
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1..
DR PROSITE; PS01186; EGF_2; UNKNOWN_1..
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match
18.1%; Score 405.5; DB 13; Length 302;

```

```

Best Local Similarity 45.2%; Pred. No. 6.7e-28;
Matches 71; Conservative 30; Mismatches 53; Indels 3; Gaps 2;

QY 11 DEETSHQLLCDKCPGGTYLKQHTAKWKTVCAPCPDHYHSDSWHTSDECLYCSPVCKELQ 70
DB 27 DRYSGLSIVCDRPPGTYLAPCSAMRKSDCAECPNAGYTFEWNHITSKLCRCS-MCAENQ 85
QY 71 YVQECNRTNHRVCECKEGRYL--ETFECLKHRSCPPGFGVQAGTPERTNTVCKRCPDGF 128
DB 86 VVQECSPSNCECKEGRYFNKKYEACIKHKECPGPGYANTTGTPHDTQTECVQCAAGF 145
QY 129 FSNETSAPCRKHTNCSVFGLLLTKQGNATHDNICS 165
DB 146 YSEVSSAKATCLAQSNCKVGLRVLLKGQDWHNTLCA 182

RESULT 8
Q62327 PRELIMINARY; PRT; 459 AA.
ID Q62327;
AC Q62327;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
RT gene.";
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.
DR HSSP; P19438; 1NCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_NGFR_1; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
FT NON_TER 1 1
FT VARIANT 87 87 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match
15.4%; Score 345; DB 11; Length 459;
Best Local Similarity 25.2%; Pred. No. 2.3e-22;
Matches 100; Conservative 53; Mismatches 164; Indels 80; Gaps 12;

QY 9 HYDEETSHQLLCDKCPGGTYLKQHTAKWKTVCAPCPDHYHSDSWHTSDECLYCSPVCKE 68
DB 31 YDORKA--QMCACRCPGQGVKHFNCNTSDTVCADCEASMTQVWNFRCLSCSSCSST 88
QY 69 LQVVKQECNRTNHRVCECKEGRYLEIEF-----CLKHRSCPPGFGVQAGTPERTNTVC 121
DB 89 DQVETRACCTQQNRVCAECAGRYCALKTHSGSCRCQMRKSKCGPGFGVASSRAPNGVLC 148
QY 122 KRCPDGFFSNETSAPCRKHTNCSVFGLLLTKQGNATHDNICSGNSESTQKV----- 174
DB 149 KACAPGTFSDTSDVCRPHRICSLAI----PGNASTDAVCAPESPTLSAIPRTLVS 204
QY 175 --DKTHTCP-----PCPAPELL-----GGPSV-----FLEPPK 200
DB 205 QPEPTRSQPLDQEPGFSQTPSILTSLSGTPIIEQSTKGGISLPIGLIVGTVSLGLMLGL 264
QY 201 PKDTLMSRPEVTVCVVVDVSHEDPEVKFNWYVDGVGVHNAKTPREEQYNSTYRVVSVL 260

```

Db 265 VNCFILVQRKKKPCQLORDA--KVPHVDEKSDQAVGL-----EQQH-----LL 306
QY 261 TVLHQDWLNGCKYCKVSNKALPA---PIEKTISKAGQPREQVYTLPPSRDELTKNOV 317
Db 307 TTPAPSSSSSSLESASAGDRAPPGGHPQARVMAEAGSQEARSRRSSDSSHGSHGTHV 366
QY 318 SLTCLVKGFYPSDIAVEWESN-----GQPNNTKTP 349
Db 367 NVTCIIVNCSSDSSQCSQASATVGDPAKPSASP 403
RESULT 9
Q16042 PRELIMINARY; PRT; 439 AA.
AC Q16042;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91370690; PubMed=1966549;
RA Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,
RA Brockhaus M., Lesslauer W.;
RT "Two human TNF receptors have similar extracellular, but distinct
RT intracellular, domain sequences."
RL Cytochrome 2:231-237(1990).
DR EMBL; S63368; AAB19824.1; .
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; .
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
DR PRODOM; PD000771; .; 1.
SQ SEQUENCE 439 AA; 46090 MW; FEBCB329CC67FF6 CRC64;

Query Match 15.3%; Score 343.5; DB 4; Length 439;
Best Local Similarity 26.8%; Pred. No. 3e-22;
Matches 107; Conservative 44; Mismatches 163; Indels 85; Gaps 15;
QY 9 HYDETSHQLLCKPCPGTYLKQHTCAKWKVCAPCPHYTDSTWHTSDECLY 68
Db 23 YDQ--TAQMCCKSCSPGQAHKVFCTKTSYVCDSCEDSTYQLWNWVPECLSCGSGRSS 80
QY 69 LQYVKQECNRTHNRVCEKGRYLEI-----EFLKHSRCPGPGVVGAGTPERTVCK 122
Db 81 DQVETQACTREQNRICTRPGWYCALSKQEGRCALPKRCRPGFGVARGPTSDVVCK 140
QY 123 RCPDGFNSSTSKAPCRKHTNCSVFGLLLTQGNATHNICSGNES--TKVDKTRCP 181
Db 141 PCAPGTFSNTSDTICRPHQICNVAI-----PGNASMDVCTSTSPRSNAPAGVHLPO 196
QY 182 PC-----PAPELLGPGSV-FLPFPKPD-----LMSRTP-----E 212
Db 197 PVSTRSQHTQPTPEFTAPSTSFLLPMGSPAPAEAGTGDFAIPVGLIVGVTLGLLIGV 256
QY 213 VTCVVVDVSHEDP-----EVKFNVDQVEVHNAK--TKPREQYNSTYRVVSVLTVLHQD 266
Db 257 VNCVIMTVQKKKPLCLQREAK-----VPHLPADKARGTGGPEQGH-----LLITAPSS 304
QY 267 WLNGKEYCKVSNKALPAPIEKTISKAGQPREQVYTLPPSRDELTK-----N 315
Db 305 SSSLESSASALDRAP-----TRNQFQAPGVEASGAGEARASTGSSDSSPGGHGT 355
QY 316 QVSLTCLVKGFYPSDIAVEWESN-----GQPNNTKTP 349
Db 356 QVNVTCIIVNCSSDSSQCSQASSTMGDTSSPSES 394

RESULT 10
O88734 PRELIMINARY; PRT; 482 AA.
ID O88734;
AC O88734;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE P80 TNF-ALPHA RECEPTOR.
GN TNFR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hurle B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;
RT "The Mouse Tumor Necrosis Factor Receptor Gene:Genomic Structure and
RT Characterization of the two Transcripts."
RL Genomics 0:0-0(0).
DR EMBL; Y14619; CAA74969.1; .
DR EMBL; Y14620; CAA74969.1; JOINED.
DR EMBL; Y14621; CAA74969.1; JOINED.
DR EMBL; Y14622; CAA74969.1; JOINED.
DR EMBL; Y14623; CAA74969.1; JOINED.
DR EMBL; Y14679; CAA74969.1; JOINED.
DR HSSP; P19438; INCF.
DR INTERPRO; IPR001368; .
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
DR PRODOM; PD000771; .; 1.
SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;
Query Match 15.1%; Score 338; DB 11; Length 482;
Best Local Similarity 25.1%; Pred. No. 1e-21;
Matches 102; Conservative 50; Mismatches 164; Indels 90; Gaps 13;
QY 9 HYDETSHQLLCKPCPGTYLKQHTCAKWKVCAPCPHYTDSTWHTSDECLY 61
Db 46 YDKRA--QMCCKACPPQYVHFCNKTSDTVCAADSDTVCADEASMTQVWNOFRTCLS 103
QY 62 CSPVCKELQYVKQECNRTHNRVCEKGRYLEIEP-----CLKHSRCPGPGVVGAGT 114
Db 104 CSSSCSTDQVETRACTKQONRVCAACEAGRYCALKTHSGRCQCMRLSKGPGFGVASSRA 163
QY 115 PERNTVCRCPDGFNSSTSKAPCRKHTNCSVFGLLLTQGNATHNICSGNESSTQKV 174
Db 164 PNGNVLCACAPGTFSDTSTSDVCRPHRICSLAI-----PGNASTDVCAPESPTLSAI 219
QY 175 -----DKTHTCP-----PCPAPELL-----GGPSV----- 194
Db 220 PRTLYVSQPETRSQPLDQEPGSPQTPSILTSLSGTPPIEQSTKGISLPTGLIVGVTS 279
QY 195 -FLPFPKPDLMISRTPEVTCVVVD--VSHEDPEVKFNWYVDGVEVHNAKTKPREQYN 251
Db 280 GLLMGLVNCVFILVORKKKPCLQORDAKVPHV-PDEKSQDAGVLEQOHLHTAPSSSSSS 338
QY 252 STYRVSVSLTVLHQDWLNGKEYKCKVSNKALPA---PIEKTISKAGQPREQVYTLPPS 308
Db 339 SL-----ESSASAGDRAPPGGHPQARVMAEAGSQEARSRRSSDSS 380
QY 309 RDELTKNOVSLTCLVKGFYPSDIAVEWESN-----GQPNNTKTP 349
Db 381 SHGSHGTHVNVTCIIVNCSSDSSQCSQASATVGDPAKPSASP 426
RESULT 11
O75509 PRELIMINARY; PRT; 655 AA.
ID O75509
AC O75509;
DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TNFR-RELATED DEATH RECEPTOR-6 (DJ181J13.1) (DR6 (TNFR-RELATED DEATH
GN RECEPTOR-6)).
DE DR6 OR DJ181J13.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Van G., Bauer J.H., Haridas V., Wang S., Liu D., Ni J., Yu G.,
RA Vincenz C., Agarwal B.B., Dixit V.M.;
RT "Identification and functional characterization of DR6, a novel death
RT domain-containing TNF receptor."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Parker A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF068868; AAC34583.1; -;
DR EMBL: AL096801; CAB75692.1; -;
DR HSSP: P07174; INGR.
DR INTERPRO: IPR000488; -;
DR INTERPRO: IPR001368; -;
DR PFAM: PF00020; TNFR_c6; 4.
DR PFAM: P00531; death; 1.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS00017; DEATH_DOMAIN; 1.
DR PROSITE: PS00050; TNFR_NGFR_2; 1.
DR PRODOM: PD000771; -; 1.
KW Receptor.
SQ SEQUENCE 655 AA; 71844 MW; 48939391C4852A33 CRC64;

Query Match 13.3%; Score 297; DB 4; Length 655;
Best Local Similarity 28.3%; Pred. No. 5.9e-18;
Matches 87; Conservative 44; Mismatches 134; Indels 42; Gaps 11;

QY 7 YLHYDETSQHLCDKCPPTYLKQHTAKWTVACPDHYTDSWHTSDECLYCSVPC 66
Db 54 YRHVDRTAGVLCDKCPAGTYVSEHCTNFTSLRVCSGCVGTTRHENGIEKCHDCSQPC 113
QY 67 KELQYVQECNRTHNRVCECKEGRYLEIEPCLKHSRCPGFGVQAGTPERNTVCKRCPD 126
Db 114 PWFIEKLPALTDRECTCPGFMFQSNATCAPHTVCPVGWGVKKGKTETEDVRCKOCAR 173
QY 127 GFPSNETSSKAPCKRKHNCVSFGLLLTKQGNATHDNICS---GNSESTQKVDKTHTCPPC 183
Db 174 GTESDVPSSVMKCAVTDCLSQLNLVVIKPGTKETDNVCGTLPFSFSSSTS----- 222
QY 184 PAPELLGGPSVFLFPKPKDTLMSRTPETVCVVVDVSHED-----PEVKFNWYVDG 235
Db 223 PSP-----GTAIF---PRP-EHMETHEVPSTYYPKGMNSTESNASSVRPKV-LSSIQEG 273
QY 236 VEVHNATKPREQYNSTYRWVSLTVLHODWLNGLNGKEYCKKVSNKALPAPI-----EKTIS 291
Db 274 TVPDNTSSARGKEDVNTK---LPNLQVNVHQ---QGPHRHIL--KLLPSMEATGGEKSS 326
QY 292 KAKGQPR 298
Db 327 PIKGPKR 333

RESULT 12
Q90544
ID Q90544 PRELIMINARY; PRT; 684 AA.
AC Q90544;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.
OS Ginglymostoma cirratum (Nurse shark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RX MEDLINE=95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks."
RL Nature 374:168-173(1995).
DR EMBL: U18701; ABA48195.1; -;
DR HSSP: P01857; 1FC1
DR INTERPRO: IPR003006; -;
DR PFAM: PF00047; 19; 6.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_3.
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 12.5%; Score 280.5; DB 13; Length 684;
Best Local Similarity 26.5%; Pred. No. 1.7e-16;
Matches 103; Conservative 53; Mismatches 148; Indels 85; Gaps 18;

QY 51 DSWHTSDECLYCS-----PVCLEQYVQECNRTHNRVCECKEGRYLEIEPCLKHSR 102
Db 315 EEWQSGVE-YTCSAKQDQSSSTPVVVKTRKARVEPTKPLRL----- 354
QY 103 CPFGFVQAGTPERNTVCKRCPDGFESNETS---SKAPCRKHTNCVSFGLLLTQ----- 154
Db 355 LPSPPEIQSTSATLTCLIR---GFYDPKVSVMQKDDVSVSANVTNFTALEQDLTFES 411
QY 155 -----KGNATHDNICSGN---SEST-----OKVDKTHTCPPCAPPELLGGP 192
Db 412 TRSLNLNLTAVENKSGAKY--TCTASHPPSQSTVKRVIRNOKVD-----CROTDI+--- 458
QY 193 SVLEFPKPKDTLMSRTPETVCVVVDVSHEDPE-VKFNWYVDGVEVHNATKPREQYN 251
Db 459 SVSLKLP-PPEEIWTQQTATIVCEIV---YSDLENIKVFWQVNGVERKKGVTQNPWSG 514
QY 252 STYRVSVLVLHODWLNGLNGKEYCKKVSNKALPAPIETISKAK-GOPREQVYTLPSRD 310
Db 515 SKSTIVSKLVMASEWSDSGTEYVCLVEDSELPTVKASIRKANYSQMHPPKVYLLHPSTD 574
QY 311 EL-TKNOVSLTCLVKGFYPSDIAVWESNGQ-PENNYKTTTPPVLDSDGSFFLYSKLTVDK 368
Db 575 EIDTENSATLMCLATNHPAEIYGVWMANDTLDSGYRTQVDSEKSGSFSFTDLRLTA 634
QY 369 SRWQQGVNFCSCVMHEALHN---HYTQKS 394
Db 635 AEWNSDITYSLVGHPSLNRLIRSTNKS 663

RESULT 13
Q9UP60
ID Q9UP60 PRELIMINARY; PRT; 384 AA.
AC Q9UP60;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SNC73 PROTEIN.
GN SNC73.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;

RT "Identification and characterization of SNC73, a gene which is down-regulated in colorectal cancer."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF067420; AAC19365.1; -

DR HSSP; P01825; 7FAB.

DR INTERPRO; IPR003006; -

DR PFAM; PF00047; 1g.3.

DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;

Query Match 12.4%; Score 277.5; DB 4; Length 384;
 Best Local Similarity 28.6%; Pred. No. 1.6e-16;
 Matches 95; Conservative 35; Mismatches 141; Indels 61; Gaps 13;

QY 115 PERNTVCKRCPDQGSNE-----TSSKAPCRKHTNCSVFGLLLTQK 155

DB 49 PGNVVIACLVQGFPOEPLSVTWSESGGVTVARNPPSQDASGLDITTSQTLTPATQ- 107

QY 156 GNATHNICSGNSESTQKVDKTHTCP-----PCPAPELLGGPSVFLFPKPK- 202

DB 108 -----CLAGKSVTCHV-KHTYNPSQDVTVPVCPVSPST-PPTSPSCCHPR 157

QY 203 -----DTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVGVHNAKTKPREEQYNSTY 254

DB 158 LSLHRLPALEDLLGSEANLTCTLTGL-RDASGVTFWTTPSSGK--SAVQPPERDLQCY 214

QY 255 RVVSVLTVLHODWLNKGYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTK 314

DB 215 SVSSVLPCCAEPNWKYFTCTAAPESTKPLTALSKS-GWTFREVHLLPPSEELAL 273

QY 315 NO-VSLTCLVKGYFSDIAVENSNGQ--PENNYKTPPVLD-SDG--SFFLYSKLTVDK 368

DB 274 NELVLTCLARGFSPKDVLRVWLOGSQELPREKYLTVASRQEPSQGTTFVAVTSILRVAA 333

QY 369 SRWQGNVFCSCVMHEALHNHYTKQSLSPG 400

DB 334 EDWKKGDTFSCWVGHEALPLAFTQKTIDRLAG 365

RESULT 14

Q9NPP6

ID Q9NPP6

AC Q9NPP6

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Pluvient R., Estivill X., Escarceller M., Sumoy L.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Auftray C., Ansoerge W., Ballabio A., Estivill X., Gibson K.,

RA Lehrach H., Poustka A., Lundeberg J.;

RT "The European IMAGE consortium for integrated Molecular analysis of

RT human gene transcripts."

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL389978; CAB97534.1; -

FT NON_TER 1

SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8A4687 CRC64;

Query Match

Best Local Similarity 12.2%; Score 273.5; DB 4; Length 416;

Matches 95; Conservative 37; Mismatches 140; Indels 69; Gaps 14;

QY 104 PGFGVQAGTPERTVCKRC-PDGFSSNE-----TSSKAPCRKHT 143

DB 82 PKVFFLSLDSTPDQGNVNVVACLVQGFPOEPLSVTWSESGGVTVARNPPSQDASGLDIT 141

QY 144 NCSVGLLLTQKGNATHNICSGNSESTQKVDKTHTCP-----PCPAPELLGGPSVFLFP 197

DB 142 TSSQTLTPATQ-----CPDGKSVTCHV-KHTYNPSQDVTVPVCPV- 180

QY 198 PPKP-----KDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVGVHNAKTKP 245

DB 181 PPPPCCHPRLSLHRLPALEDLLGSEANLTCTLTGL-RDASGVTFWTTPSSGK--SAVQPP 237

QY 246 REEQNYSYRVSVLTVLHODWLNKGYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 305

DB 238 PERDLQCYSVSSVLPCCAEPNWKYFTCTAAPESTKPLTALSKS-GWTFREVHLLPPSE 296

QY 306 PPSRDELTKNQ-VSLTCLVKGYFSDIAVENSNGQ--PENNYKTPPVLD-SDG--SFF 359

DB 297 PPSSEALNELVLTCLARGFSPKDVLRVWLOGSQELPREKYLTVASRQEPSQGTTFV 356

QY 360 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPG 400

DB 357 VTSILRVAAEDWKKGDTFSCWVGHEALPLAFTQKTIDRLAG 397

RESULT 15

Q9Y6Q6

ID Q9Y6Q6

AC Q9Y6Q6

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE RECEPTOR ACTIVATOR OF NF-KAPPA-B PRECURSOR (TNF-RELATED ACTIVATION-INDUCED CYTOKINE RECEPTOR) (RANK).

GN TNFRSF11A OR RANK.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98032977; PubMed=9367155;

RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,

RA Tometsko M.E., Roux E.R., Teepe M.C., DuBoise R.F., Cosman D.,

RA Galibert L.;

RT "A homologue of the TNF receptor and its ligand enhance T-cell growth

RT and dendritic-cell function."

RL Nature 390:175-179(1997).

CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS RANKL.

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).

CC -!- TISSUE SPECIFICITY: UBIQUITOUS EXPRESSION WITH HIGH LEVELS IN

CC SKELETAL MUSCLE, THYMUS, LIVER, COLON, SMALL INTESTINE AND ADRENAL

CC GLAND.

CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

DR EMBL; AF018253; AAB86809.1; -

DR HSSP; P25942; ICDF.

DR MIN; 603499; -

DR INTERPRO; IPR001368; -

DR PFAM; PF00020; TNFR_C6; 4.

DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

DR PROSITE; PS50050; TNFR_NGFR_2; 1.

KW Receptor; Glycoprotein; Transmembrane; Signal; Repeat.

FT SIGNAL 1 22

FT CHAIN 23 616

FT DOMAIN 24 212

FT TRANSMEM 213 233

FT DOMAIN 234 616

FT DOMAIN 33 195

FT REPEAT 33 69

FT REPEAT 70 112

FT REPEAT 113 152

FT REPEAT 153 195

FT DISULFID 34 46

FT DISULFID 47 60

BY SIMILARITY.

```

FT DISULFID 50 68 BY SIMILARITY.
FT DISULFID 71 86 BY SIMILARITY.
FT DISULFID 92 112 BY SIMILARITY.
FT DISULFID 114 124 BY SIMILARITY.
FT DISULFID 126 133 BY SIMILARITY.
FT DISULFID 131 151 BY SIMILARITY.
FT DISULFID 127 169 BY SIMILARITY.
FT DISULFID 154 194 BY SIMILARITY.
FT DISULFID 175 194 BY SIMILARITY.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 616 AA; 66033 MW; E3DE9A7A08196F81 CRC64;

Query Match 11.5%; Score 258.5; DB 4; Length 616;
Best Local Similarity 34.5%; Pred. No. 1.3e-14;
Matches 58; Conservative 21; Mismatches 82; Indels 7; Gaps 4;

QY 20 CDKCPGTYLKQCHTAKWTVACPDHYTDSWHTSDECLYCPVC---KELOYVKQEC 76
Db 47 CNKCEFGKYMSSKCTTTSDSVCLPCGPDYLDVSWNEEDKCL-LHKVCDTGKALVAVVAG- 104

QY 77 NRTHNRVCECKEGRY--LEIEFCLKHRSCPPGFGVVQAGTPERTVCKRCPDGFSSNETS 134
Db 105 NSTTPRRCACTAGYHWSQDCECCRRNTECAPGLGAQHPLQLNKDVTCKPCLAGYFSDAFS 164

QY 135 SKAPCRKHTNCSVFGLLLTQKGNATHDNCISGENSESTOKVDKTHTCPP 182
Db 165 STDKCRPWNTCTFLGKRVEHHGTEKSDAVCSSSLPARKPPNEPHVYLP 212

```

Search completed: March 1, 2001, 09:20:07
Job time: 405 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:44 ; Search time 40.97 seconds
(without alignments)
315.295 Million cell updates/sec

Title: US-09-389-782A-5
Perfect score: 2240
Sequence: 1 ETPPKYLHYDEETSHQLC.....VMHEALHNHYTKSLSPG 400

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1234	55.1	330	1 GCL_HUMAN	P01857 homo sapien
2	1148.5	51.3	326	1 GC2_HUMAN	P01859 homo sapien
3	1136	50.7	327	1 GC4_HUMAN	P01861 homo sapien
4	1133.5	50.6	320	1 GC3_HUMAN	P01860 homo sapien
5	923	41.2	323	1 GC_RABIT	P01870 oryctolagus
6	904.5	40.4	329	1 GC2_CAVPO	P01862 cavia porce
7	847.5	37.8	329	1 GC3_MOUSE	P22436 mus musculus
8	843	37.6	333	1 GCB_RAT	P20761 rattus norv
9	841.5	37.6	398	1 GC3M_MOUSE	P03987 mus musculus
10	820	36.6	324	1 GCL_MOUSE	P01868 mus musculus
11	820	36.6	326	1 GCL_RAT	P20759 rattus norv
12	820	36.6	393	1 GC1M_MOUSE	P01869 mus musculus
13	812.5	36.3	330	1 GCRA_MOUSE	P01863 mus musculus
14	812.5	36.3	399	1 GCAM_MOUSE	P01865 mus musculus
15	804.5	35.9	329	1 GCC_RAT	P20762 rattus norv
16	801.5	35.8	335	1 GCAB_MOUSE	P01864 mus musculus
17	791	35.3	322	1 GCA_RAT	P20760 rattus norv
18	777.5	34.7	336	1 GCB_MOUSE	P01866 mus musculus
19	777.5	34.7	405	1 GCBM_MOUSE	P01867 mus musculus
20	365.5	16.3	391	1 MUCB_HUMAN	P04220 homo sapien
21	361	16.1	421	1 EPC_MOUSE	P06336 mus musculus
22	360	16.1	454	1 MUC_HUMAN	P01871 homo sapien
23	354.5	15.8	455	1 MUC_MOUSE	P01872 mus musculus
24	353	15.8	428	1 EPC_HUMAN	P01854 homo sapien
25	353	15.8	429	1 EPC_RAT	P01855 rattus norv
26	349.5	15.6	476	1 MUCM_MOUSE	P01873 mus musculus
27	345.5	15.4	461	1 TNR2_HUMAN	P20333 homo sapien
28	340	15.2	458	1 MUC_RABIT	P03988 oryctolagus
29	340	15.2	474	1 TNR2_MOUSE	P25119 mus musculus
30	338.5	15.1	450	1 MUC_CANFA	P01874 canis fami
31	335	15.0	479	1 MUCM_RABIT	P04221 oryctolagus
32	333.5	14.9	454	1 MUC_MESAU	P06337 mesocricetu
33	332.5	14.8	457	1 MUC_SUNNU	P20768 suncus muri

ALIGNMENTS

RESULT 1

ID	GCL_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG GAMMA-1 CHAIN C REGION.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-82274238; PubMed-6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF I-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE-71064024; PubMed-5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
RN	[3]			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE-71064025; PubMed-5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
RN	[4]			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE-77070269; PubMed-826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein Nie) III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE-83289131; PubMed-6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS			
RX	MEDLINE-71064027; PubMed-4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X. Intrachain disulfide bonds."			

P23085 heterodontu
P23087 heterodontu
P01879 oryctolagus
P23088 heterodontu
P23086 heterodontu
P27512 mus musculu
P01875 gallus gall
P23084 heterodontu
P01876 homo sapien
P20758 gorilla gor
P01877 homo sapien
P25942 homo sapien

34 315.5 14.1 438 1 HVC2_HETFR
35 299 13.3 438 1 HVC3_HETFR
36 298 13.3 299 1 ALC_RABIT
37 295 13.2 461 1 HVCN_HETFR
38 288 12.9 393 1 HVC3_HETFR
39 287.5 12.8 289 1 CD40_MOUSE
40 282.5 12.6 446 1 MUC_CHICK
41 279.5 12.5 370 1 HVC1_HETFR
42 277.5 12.4 353 1 ALC1_HUMAN
43 274.5 12.3 353 1 ALC1_GORGO
44 273.5 12.2 340 1 ALC2_HUMAN
45 264 11.8 277 1 CD40_HUMAN

RL Biochemistry 9:3188-3196(1970).
RN [7]
RP DISULFIDE BONDS.
RX MEDLINE=770267; PubMed=1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT "Rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
RT characterization of the protein, the L- and H-chains, the
RT cyanogen bromide cleavage products, and the disulfide bridges";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=81208100; PubMed=7236608;
RA Deisenhofer J.;
RT "Crystallographic refinement and atomic models of a human Fc fragment
RT and its complex with fragment B of protein A from Staphylococcus
RT aureus at 2.9- and 2.8-A resolution.";
RL Biochemistry 20:2361-2370(1981).
CC -!- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE
CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)
CC MARKER & THE GIM (NON-1) MARKERS.
CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
CC 35,116,198,269 & 272.
CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
CC 268-272.
CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
CC RESIDUES 198,267&272.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00228; AAC82527.1; ALT_INIT.
DR PIR; A02146; GHU.
DR PDB; 1FC1; 15-JUL-92.
DR PDB; 1FC2; 15-JUL-92.
DR MIM; 147100; -.
DR INTERPRO; IPR000495; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 223 CH2.
FT DOMAIN 224 330 CH3.
FT DISULFID 27 83
FT DISULFID 103 103
FT DISULFID 109 109 INTERCHAIN (WITH LIGHT CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 144 204 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 250 308
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .).
FT VARIANT 97 97 /FTid=VAR_003886.
FT VARIANT 239 239 D -> E (IN GIM(NON-1) MARKER).
FT VARIANT 241 241 /FTid=VAR_003887.
FT MOD_RES 330 330 L -> M (IN GIM(NON-1) MARKER).
FT STRAND 123 126 /FTid=VAR_003888.
FT HELIX 130 134 REMOVED POST-TRANSLATIONALLY.
FT TURN 136 137
FT STRAND 141 148
FT STRAND 158 162
FT TURN 163 164

FT STRAND 165 166
FT STRAND 175 178
FT STRAND 183 190
FT HELIX 193 197
FT TURN 198 199
FT STRAND 202 206
FT STRAND 215 219
FT STRAND 227 227
FT STRAND 230 234
FT STRAND 238 240
FT TURN 241 242
FT STRAND 245 256
FT STRAND 260 266
FT TURN 267 268
FT STRAND 269 270
FT STRAND 274 276
FT STRAND 280 281
FT TURN 283 284
FT STRAND 287 296
FT HELIX 297 301
FT TURN 302 303
FT STRAND 306 312
FT TURN 313 314
FT TURN 316 317
FT STRAND 320 324
SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;

Query Match 55.1%; Score 1234; DB 1; Length 330;
Best Local Similarity 92.1%; Pred. No. 6.1e-77;
Matches 232; Conservative 4; Mismatches 6; Indels 10; Gaps 2;

QY 159 THDNICS-----GNSESQKV-----DKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMIS 208
DB 78 TQYICNVNHPSTKVDKVPKSCDTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMIS 137
QY 209 RTPVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVSVLTVQLQDWL 268
DB 138 RTPVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVSVLTVQLQDWL 197
QY 269 NGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGY 328
DB 198 NGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGY 257
QY 329 SDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSCVMHEALHN 388
DB 258 SDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSCVMHEALHN 317
QY 389 HVTQKSLSLSPG 400
DB 318 HVTQKSLSLSPG 329

RESULT 2
GC2_HUMAN
ID GC2_HUMAN STANDARD; PRT; 326 AA.
AC P01859;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2 CHAIN C REGION.
GN IGHG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82197621; PubMed=6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
RT heavy chain constant region genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
RN [2]


```

FT DISULFID      109          109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID     140          200
FT DISULFID     246          304 S -> A (IN MYELOMA PROTEINS TIL & ZIE).
FT VARIANT       60           60 /FTID-VAR_003889.
FT SITE         156          156 AT OR NEAR THE COMPLEMENT-BINDING SITE..
FT MOD_RES      326          326 REMOVED POST-TRANSLATIONALLY (PROBABLE)..
FT SEQUENCE     326 AA; 35884 MW; 8310878C6878CF9C CRC64;

Query Match              51.3%; Score 1148.5; DB 1; Length 326;
Best Local Similarity   90.3%; Pred. No. 3.7e-71;
Matches 214; Conservative 8; Mismatches 8; Indels 7; Gaps 2;

QY 170 STQVDKT-----HTCPPCPAPALLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDYSHE 223
DB I II III : : : : : : : : : : : : : : : : : : : : : : : : : : :
90 SNTKVDKTVRKCCVECPCEAPP-VAGPSVFLPFPKPKDTLMISRTPEVTCVVVDYSHE 148

QY 224 DPEVKFNWYDGVGVHNAKTKPREEQNSTYRVVSIVTLVHQDWLNGKEYCKYKSVNKAALP 283
DB I II III : : : : : : : : : : : : : : : : : : : : : : : : : : :
149 DPEQVFNWYDGVGVHNAKTKPREEQNSTFRVSVLTVVHQDWLNGKEYCKYKSVNKGLP 208

QY 284 APIEKTISKAGQPREPOVYITPLPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGOPEN 343
DB I II III : : : : : : : : : : : : : : : : : : : : : : : : : : :
209 APIEKTISKTKGPREPOVYITPLPSREMTKNQVSLTCLVKGFYPSDIAVEMESNGOPEN 268

QY 344 NYKTTPTVLDSGFFLYSKLTDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG 400
DB I II III : : : : : : : : : : : : : : : : : : : : : : : : : : :
269 NYKTTPTMLDSGFFLYSKLTDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG 325

RESULT 3
GC4_HUMAN
ID GC4_HUMAN STANDARD; PRT; 327 AA.
AC P01861;
DT 21-JUL-1986 (Rel. 01; Created)
DI 21-JUL-1986 (Rel. 01; Last sequence update)
DE 15-JUL-1999 (Rel. 38; Last annotation update)
DE IG GAMMA-4 CHAIN C REGION.
GN IGHG4
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
SEQUENCE FROM N.A.
MEDLINE=83157104; PubMed=6299662;
RA Ellison J.W., Buxbaum J.N., Hood L.E.;
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL DNA 1:11-18(1981).
[2]
SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE=70207560; PubMed=4192699;
RA Plunk J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
RT "Human Immunoglobulin subclasses. Partial amino acid sequence of the constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
CC EMBL; K01316; RA059394.1; ALT_INIT.
DR PIR; A02150; G4HU.
DR MIG; 147130; -.
DR INTERPRO; IPR000495; -.
DR INTERPRO; IPR003006; -.
DR DR PFAM; PF000047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
```


SQ SEQUENCE 290 AA; 32331 MW; B69CBC95705B2F46 CRC64;
 Query Match 50.6%; Score 1133.5; DB 1; Length 290;
 Best Local Similarity 70.88; Pred. No. 3.4e-70;
 Matches 218; Conservative 19; Mismatches 22; Indels 49; Gaps 5;
 QY 97 CLKHRS-PPGFGVVQAGTPTERTVCKRCPCDGFNETSSKAP-CRKHTNCSVFGLL 152
 Db 27 CPEPKSCDTTPP-----CPCPEP-KSCDTPPCPCPEPKSC----- 63
 QY 153 TOKGNATHNICSNSESTOKVTHRCPCPCAPBELLGGPSVFLPPKPKDTLMISRTPE 212
 Db 64 -----DTPPCPCPCAPBELLGGPSVFLPPKPKDTLMISRTPE 101
 QY 213 VTCVVVDVSHEDPVEKFNWYVGVGEVHNATKPREQYNSYRVSVLTVLHQQDLNGKE 272
 Db 102 VTCVVVDVSHEDPVEKFNWYVGVGEVHNATKPREQYNSYRVSVLTVLHQQDLNGKE 161
 QY 273 YKCVSNKALPAPIEKTISKAKGQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIA 332
 Db 162 YKCVSNKALPAPIEKTISKAKGQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIA 221
 QY 333 VEWESGQPNNTYTPPVLDSGDFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQ 392
 Db 222 VEWESGQPNNTYTPPVLDSGDFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQ 281
 QY 393 KSLSLSPG 400
 Db 282 KSLSLSPG 289
 RESULT 5
 GC_RABIT STANDARD; PRT; 323 AA.
 AC POL1870;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA CHAIN C REGION.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=84030930; PubMed=6313520;
 RA Bernstein K.E., Alexander C.B., Mage R.G.;
 RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
 F-1 haplotype.";
 RL Immunogenetics 18:387-397(1983).
 RN [2]
 RN SEQUENCE OF 1-128.
 RX MEDLINE=76135469; PubMed=1243651;
 RA Pratt D.M., Mole L.E.;
 RT "Sequence studies on the constant region of the Fd sections of rabbit
 immunoglobulin G of different allotype.";
 RL Biochem. J. 151:337-349(1975).
 RN [3]
 RN SEQUENCE OF 88-266 FROM N.A.
 RX MEDLINE=83299917; PubMed=6193512;
 RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
 RT "Heavy chain genes of rabbit IgG: Isolation of a cDNA encoding gamma
 heavy chain and identification of two genomic C gamma genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
 RN [4]
 RN SEQUENCE OF 132-161.
 RX MEDLINE=70110015; PubMed=5461106;
 RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
 RT "Sequence studies of the Fd section of the heavy chain of rabbit
 immunoglobulin G.";
 RL Biochem. J. 116:249-259(1970).
 RN [5]
 RN SEQUENCE OF 129-131 AND 155-322.

RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
 RL (In) Killander J. (eds.);
 RL Gamma globulins, Nobel symp. 3, pp.109-127, Almquist and Wiksell,
 Stockholm (1967).
 CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
 CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
 CC MARKERS AND REF.5 THE E15 MARKER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M16426; AAA31289.1; -;
 DR PIR; A02161; GHRB.
 DR INTERPRO; IPR000495; -;
 DR INTERPRO; IPR003006; -;
 DR PFAM; PF00047; Ig; 3.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1
 FT VARIANT 104 104 T -> M (IN D11 MARKER).
 FT VARIANT 185 185 T -> A (IN E15 MARKER).
 FT CONFLICT 48 48 N -> E (IN REF. 2).
 FT CONFLICT 71 71 V -> VPV (IN REF. 2).
 FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
 FT CONFLICT 173 173 N -> D (IN REF. 5).
 FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
 FT CONFLICT 201 201 N -> D (IN REF. 5).
 FT CONFLICT 218 218 Q -> E (IN REF. 5).
 FT CONFLICT 233 233 E -> Q (IN REF. 5).
 FT CONFLICT 246 246 N -> D (IN REF. 5).
 FT CONFLICT 256 256 E -> G (IN REF. 5).
 FT CONFLICT 260 260 N -> D (IN REF. 5).
 FT CONFLICT 266 266 N -> D (IN REF. 5).
 FT CONFLICT 280 280 Y -> W (IN REF. 5).
 FT CONFLICT 284 284 N -> S (IN REF. 5).
 SQ SEQUENCE 323 AA; 35404 MW; 69E8AAL18D579AB8 CRC64;
 Query Match 41.2%; Score 923; DB 1; Length 323;
 Best Local Similarity 61.58; Pred. No. 6.7e-56;
 Matches 177; Conservative 33; Mismatches 46; Indels 32; Gaps 5;
 QY 123 RCPDGFES-NTSSKAPCRKHTNCSVFGLLLTQKGNATHNICSNSESTOKVDTK 177
 Db 57 RQSSGLYSLSVSVSSSQP-----VTCNV-----AHPATNFKVDKT 94
 QY 178 ---HTC-PPCPAPBELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPVEKFNWY 232
 Db 95 VAPSTCSKPTCPBELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPVEKFNWY 154
 QY 233 VDGVEVHNATKPREQYNSYRVSVLTVLHQQDLNGKEKCKVSNKALPAPIEKTISK 292
 Db 155 INNEQVTRARPLREQQNFNSIRVVSTLPITHQDLWLRKEFKCKVHNKALPAPIEKTISK 214
 QY 293 AKGQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESGQPNNTYTPPVL 352
 Db 215 ARGQPLEPKVYTMGPPREELSSRSVSLTCMNGFYPSDISVEWEKNGKAEDNYKTPAVL 274
 QY 353 DSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSLSPG 400
 Db 275 DSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSLSPG 322
 RESULT 6
 GC2_CAVPO
 ID GC2_CAVPO STANDARD; PRT; 329 AA.
 AC P01862;
 DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2 CHAIN C REGION.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshstein B.K., Hussain Q.Z., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the
RT half-cysteine joining heavy and light chains.";
RL Biochemistry 10:18-25(1971).
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
RT and hinge region cyanogen bromide fragments.";
RL Biochemistry 10:9-17(1971).
RN [4]
RP SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E., Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4796-4803(1974).
RN [5]
RP SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;
RA Trischmann T.M., Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4804-4811(1974).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B., Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
RL Biochemistry 10:26-31(1971).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR; A02151; G2GP.
DR INTERPRO; IPR000495; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 16 16
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 36074 MW; 50231B7164DIFBA9 CRC64;

Query Match 40.4%; Score 904.5; DB 1; Length 329;
Best Local Similarity 60.3%; Pred. No. 1.2e-54;
Matches 173; Conservative 30; Mismatches 53; Indels 31; Gaps 3;
QY 127 GFSSNETSSKAPCRKHTNCVFGLLLTQGNATHDNCISGNSSTQKVDKT----- 177
DB 61 GLYSLTSMVTVPSSQKATCNV-----AHPASSTKVDKTVETIRTPZP 102

QY 178 --HTCPCPAPPELLGGPSVFLFPPPKPDKTLMISRTPEVTCVVVDVSHEDPEVKFNWYDG 235
DB 103 BCTCPKCPPEPPELGGPSVFLFPPPKPDKTLMISRTPEVTCVVVDVSHEDPEVKFNWYDG 162
QY 236 VEVHNAKTPREEQYNSYRVSVTLVHQLDNLNGKEYCKYKSNKALPAPIEKTISKAKG 295
DB 163 KPVGNAETKPRVEQNTTFRVESVLPVQHODWLRGKFKCKYKSNKALPAPIEKTISKAKG 222
QY 296 QPREQVYTLPPSRDELTKNQVSLTCLVKYGFPYPSDIAVWESNGQP--ENNTYKTTTPPVLD 353
DB 223 APRMPDVYTLPPSRDELTKNQVSLTCLVKYGFPYPSDIAVWESNGQP--ENNTYKTTTPPVLD 282
QY 354 SDGSFELYSLKTLVDKSRWQGVNFSCSVNHEALHNHYTQKSLSLSPG 400
DB 283 ADGSFELYSLKTLVDKSRWQGVNFSCSVNHEALHNHYTQKSLSLSPG 329
RESULT 7
GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, SECRETED FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; -; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR INTERPRO; IPR000495; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin C region; Glycoprotein; Transmembrane;
KW Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 37.8%; Score 847.5; DB 1; Length 329;
Best Local Similarity 54.0%; Pred. No. 8.9e-51;
Matches 162; Conservative 47; Mismatches 72; Indels 19; Gaps 5;
QY 107 FGWVQAGTPERTVCKPCDGFSSNETSSKAPCRKHTNCVFGLLLTQGNATHDNCISG 166
DB 42 YGALSSGV---RTVSSVLQSGFYSLSLVTPVSSWPSQTVI-----CNVAHP---AS 88
QY 167 NSESTQKVD---KTHTCP--PCPAPPELLGGPSVFLFPPPKPDKTLMISRTPEVTCVVVDV 220
DB 89 KTELKRIEPRIPKPTSPGSSCPGNILGGPSVFIFFPKPKDALMISLTPKVTCCVVVDV 148


```

FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 102 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 36.6%; Score 820; DB 1; Length 326;
Best Local Similarity 60.5%; Pred. No. 6.4e-49;
Matches 146; Conservative 42; Mismatches 39; Indels 14; Gaps 3;

QY 170 STOKVDKTH-----HTCRPPAPPELLGG--PSVLFPPPKDMLISRTPEVTCVVVD 219
Db 89 SSTKVDKKIVPRNCGGDKPC----ICTGEVSVFFPPKPDVLTITLTPKTCVVVD 144

QY 220 VSHEDPEVFNWYDGVVEVHNATKPREQNSTYRVSVTLVHODWLNKKEYCKVSN 279
Db 145 ISQDDPEVHFSWFVDDVEVHTAOTRPEEQFNSTFRSVSELPILHODWLNKGRTRCKVTS 204

QY 280 KALPAIEKTIKAKGPREPQVYITPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNG 339
Db 205 AAFPSPIEKTIKSGRQVPHVYITMSPTKEEMTQNEVSTCMVKGFYPPDIYVQWQNG 264

QY 340 OPENNYKTPPVLDSDGSFELYSLKLVDRKRWQGVNFCVSMHEALHNHYTKSLSP 399
Db 265 QPENYKTPPVLDSDGSFELYSLKLVDRKRWQGVNFCVSMHEALHNHYTKSLSP 324

QY 400 G 400
Db 325 G 325

RESULT 12
ID GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowan A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.

```

```

RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
SEGMENT OF MU CHAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licensed@isb-sib.ch).
CC EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02159; GIMSM.
DR MGI; MGI:96446; IGH-4.
DR INTERPRO; IPR000495; -
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; Igh; 3.
DR PROSITE; PS00290; IGH_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 HINGE.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC...).
FT DISULFID 244 302
FT TRANSMEM 340 357 POTENTIAL.
FT DOMAIN 358 393 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 36.6%; Score 820; DB 1; Length 393;
Best Local Similarity 60.5%; Pred. No. 7.9e-49;
Matches 144; Conservative 45; Mismatches 39; Indels 10; Gaps 3;

QY 170 STOKVDKTH-----CPP--CPAPELLGGPSVLFPPPKDMLISRTPEVTCVVVDVSH 222
Db 89 SSTKVDKKIVPRDCGCKPCICTVPEV---SSVFPPPKPDVLTITLTPKTCVVVDISK 145

QY 223 EDEVEFNWYDGVVEVHNATKPREQNSTYRVSVTLVHODWLNKKEYCKVSNKAL 282
Db 146 DDEVEQFSWFVDDVEVHTAOTRPEEQFNSTFRSVSELPIMHODWLNKKEYCKRVNSAAF 205

QY 283 PAPEKTIKAKGPREPQVYITPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPE 342
Db 206 PAPEKTIKAKGPREPQVYITPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPE 265

QY 343 NNYKTPPVLDSDGSFELYSLKLVDRKRWQGVNFCVSMHEALHNHYTKSLSPG 400
Db 266 ENYKTPPVLDSDGSFELYSLKLVDRKRWQGVNFCVSMHEALHNHYTKSLSPG 323

RESULT 13
GCAA_MOUSE STANDARD; PRT; 330 AA.
ID GCAA_MOUSE

```

AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RT Balb/c gamma 2a heavy chain messenger RNA.";
RL Nucleic Acids Res. 8:3143-3155(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6267279;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
RT and evolution of heavy chain genes: further evidence for intervening
RT sequence-mediated domain transfer.";
RL Nucleic Acids Res. 9:1365-1381(1981).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Ollo R., Auffray C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
RT suggests that exons can be exchanged between genes in a multigenic
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN [4]
RP MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin: amino-acid sequence of the Fc fragment. Implications
RT for the evolution of immunoglobulin structure and function.";
RL Eur. J. Biochem. 43:423-435(1974).
RN [5]
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
RT immunoglobulin. Identification of the disulfide bridges.";
RL Eur. J. Biochem. 30:452-462(1972).
RN [6]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V00798; CAA24178.1; -
CC PIR; A02152; G2MSA.
CC INTERPRO; IPR000495; -
CC INTERPRO; IPR003006; -
CC PFAM; PF000047; Ig; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region.
CC NON_TER 1 15
CC DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
CC DISULFID 27 82
CC DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 144 204
CC DISULFID 250 308
CC DISULFID 330 330 REMOVED POST-TRANSLATIONALLY.
CC SEQUENCE 330 AA; 36389 MW; B84361C545A6864 CRC64;

Query Match 36.3%; Score 812.5; DB 1; Length 330;
Best Local Similarity 56.5%; Pred. No. 2.le-48;
Matches 157; Conservative 34; Mismatches 58; Indels 29; Gaps 4;
Qy 133 TSSKAPCRKHTNCSVFGLLLTQGNATHDNCISGNSESTQKVDK-----THTCPP-- 182
Db 71 TSSWPSQSI-CNV-----AHPASSTKVDDKIEPRGPTIKPCPPCK 111
Qy 183 CPAPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNK 242
Db 112 CPAPNLLGGPSVFLPPPKIKDVLMSLSPITCVVVDVSEDDPDVQISWFVNNVEVHTAQ 171
Qy 243 TKPREQYNSYRYVSVLTVLHQDLNGCKYKCKVSNKALPAPIETKTSKAGGPREPOV 302
Db 172 TQTHREDYNSYRVSALPIQHDWMSGKFKCKVNNKDLPAPIERTISKPGSVRAPOV 231
Qy 303 YTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVPSGNGPKNNYKTPPVLDSDGSFFLYS 362
Db 232 YVLPPEPEETKKQVLTTCMVTDMPEDYVVENTNGKTELNYKNTPEVLDSDGSFYFWS 291
Qy 363 KLTVDKSRWQGNVFSQVSMHEALHNHYTQKSLSLSPG 400
Db 292 KLRVEKKNWVERNSYSCSVVHGLHNHTTKFSRTPG 329
RESULT 14
GCAM_MOUSE STANDARD; PRT; 399 AA.
ID GCAM_MOUSE AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGS CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J00471; AAB59661.1; ALT_INIT.
CC PIR; A02154; G2MSAM.
CC MGD; MGI:96443; IGH-1.
CC INTERPRO; IPR000495; -
CC INTERPRO; IPR003006; -
CC PFAM; PF000047; Ig; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
CC Transmembrane; Alternative splicing.
CC NON_TER 1 15
CC DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).


```
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399
FT CARBOHYD 180 180
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;

Query Match 36.3%; Score 812.5; DB 1; Length 399;
Best Local Similarity 56.5%; Pred. No. 2.6e-48;
Matches 157; Conservative 34; Mismatches 58; Indels 29; Gaps 4;

QY 133 TSSKAPCRKHTNCSVFLPGLLTQKGNATHDNCNSSESTOKVDK-----RHTCSP-- 182
Db 71 TSSTWPSQSIT-CNV-----AHPASSTKVDKKIEPRGPTIKPCPPCK 111
QY 183 CPAPELLGGPSVFLPFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 242
Db 112 CPAPNLLGGPSVFLPFPKIKDVLISLPIVTCVVVDVSEDDPDVQISWFWNNVEVHTAQ 171
QY 243 TKPREOYNSTYRVSVLVTLVHODWLNKGYCKVSKNKPAPLEKTIKAKGQPREPQV 302
Db 172 TOTREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNKDLPAPIERTISKPGSVRAPQV 231
QY 303 YTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYS 362
Db 232 YVLPFPEEEMTKQVLTCTMTVDMPEDIVYEWNTNNGKTELNYKNTPEVLDSDGSYFMS 291
QY 363 KLVVDKSRVQOQGNVSCSVMHALHNHYTKLSLSPG 400
Db 292 KLAVEKKKNWERNYSYSCSVVHGLHNHHTKFSRTPG 329

RESULT 15
GCC_RAT GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2C CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88166903; PubMed-312722;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RL region cDNA: extensive homology to mouse gamma 3."
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07189; CAA30169.1; -
DR PIR; S00847; S00847.
DR INTERPRO; IPR000495; -
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
```

```
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 35.9%; Score 804.5; DB 1; Length 329;
Best Local Similarity 63.7%; Pred. No. 7.2e-48;
Matches 142; Conservative 41; Mismatches 37; Indels 3; Gaps 1;

QY 181 PP---CPAPELLGGPSVFLPFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 237
Db 106 PPTDICSCDDNLGRPSVFIFFPKPKDILMITLTPKVTCTVVVDVSEEPDQVFSWFDNVR 165
QY 238 VHNAKTKPREOYNSTYRVSVLVTLVHODWLNKGYCKVSKNKPAPLEKTIKAKGQ 297
Db 166 VFTAQTPHEQLNGTFRVWSTLHIQHDWMSGKEFKCKVNNKDLPSPIETISKPRKA 225
QY 298 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGS 357
Db 226 RTPQVYTIPTPREQMSKNKYSLTCMTVTSFYFASISVEWERNGELEQDYKNTLPVLDSDS 285
QY 358 FFLYSKLTVDKSRVQOQGNVSCSVMHALHNHYTKLSLSPG 400
Db 286 YFLYSKLSLVDTSWNRGDIYTCVSVHVALHNHHTQKNLSRSPG 328
```

Search completed: March 1, 2001, 09:17:45
Job time: 263 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:45 ; Search time 79.26 Seconds
(without alignments)
342.673 Million cell updates/sec

Title: US-09-389-782A-5
Perfect score: 2240
Sequence: 1 ETFFPKYLHYDEETSHQLLC.....VMHEALHNHYTKSLSPG 400

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results, predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1234	55.1	330	1 GHU	Ig gamma-1 chain C
2	1230	54.9	374	2 S72664	Ig heavy chain V r
3	1229	54.9	255	4 S31866	Ig gamma-1 chain C
4	1225	54.7	374	2 S69339	Ig heavy chain V r
5	1183.5	52.8	234	2 PT0207	Ig gamma chain C r
6	1154.5	51.5	377	2 A23511	Ig gamma-3 chain C
7	1152.5	51.5	377	2 A60764	Ig gamma-3 chain C
8	1148.5	51.3	326	1 G2HU	Ig gamma-2 chain C
9	1136	50.7	327	1 G4HU	Ig gamma-2 chain C
10	1133.5	50.6	289	1 G3H0W1	Ig gamma-4 chain C
11	923	41.2	323	1 GHRB	Ig gamma-3 heavy C
12	917.5	41.0	328	2 I47160	Ig gamma chain C r
13	917.5	41.0	328	2 I47159	Ig gamma 2b chain
14	915.5	40.9	277	2 I47162	Ig gamma 2a chain
15	904.5	40.4	329	1 G2GP	Ig gamma 4 chain c
16	895.5	40.0	328	2 I47158	Ig gamma-2 chain C
17	889.5	39.7	328	2 I47161	Ig gamma 1 chain c
18	861.5	38.5	470	2 S22080	Ig heavy chain pre
19	851.5	38.0	308	2 C30554	Ig heavy chain C r
20	851.5	38.0	472	2 S31459	Ig gamma-1 chain -
21	847.5	37.8	329	1 G3MSC	Ig gamma-3 chain C
22	843	37.6	333	2 PS0018	Ig gamma-2b chain
23	841.5	37.6	398	1 G3MSM	Ig gamma-3 chain C
24	830	37.1	444	2 PC4436	monoclonal antibody
25	820	36.6	324	1 GLMS	Ig gamma-1 chain C
26	820	36.6	326	2 PS0017	Ig gamma-1 chain C
27	820	36.6	393	1 G1MSM	Ig gamma-1 chain C
28	812.5	36.3	330	1 G2MSA	Ig gamma-2a chain
29	812.5	36.3	399	1 G2MSAM	Ig gamma-2a chain

ALIGNMENTS

RESULT 1

GHU

Ig gamma-1 chain C region - human

C;Species: Homo sapiens (man)

C;Date: 31-Jan-1981 #sequence.revision 18-Aug-1982 #text.change 16-Jul-1999

C;Accession: A93433; S33887; B90563; A90564; B91668; A91723; A02146

R;Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A;Title: The nucleotide sequence of a human immunoglobulin C-gamma gene.

A;Reference number: A93433; MUID:82274238

A;Accession: A93433

A;Molecule type: DNA

A;Residues: 1-330 <ELL>

A;Cross-references: EMBL:Z17370

A;Note: this sequence has the Gim(17) allotypic marker, 97-Lys, and the Gim(1) marker

A;Note: Lys-330 is removed after translation

R;Harris, L.J.

submitted to the EMBL Data Library, October 1992

A;Reference number: S33904

A;Accession: S36861

A;Molecule type: DNA

A;Residues: 2-330 <HAR>

A;Cross-references: EMBL:Z17370

R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of

A;Reference number: S33887; MUID:83001943

A;Accession: S33887

A;Molecule type: DNA

A;Residues: 88-113;235-330 <TAK>

A;Cross-references: EMBL:Z17370

R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman

Biochemistry 9, 3161-3170, 1970

A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq

A;Reference number: A90563; MUID:71064024

A;Contents: myeloma protein Eu

A;Accession: B90563

A;Molecule type: protein

A;Residues: 1-96,'R',98-135 <CUN>

A;Note: this sequence has the Gim(3) marker, 97-Arg

R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se

A;Reference number: A90564; MUID:71064025

A;Contents: Eu

A;Accession: A90564

A;Molecule type: protein

A;Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2

A;Note: this sequence has the Gim(non-1) markers, 239-Glu and 241-Met

R;Ponstingl, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein NI

Db 145 SKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP 204
QY 351 VLSDSGSFFLYSKLTVDKSRWQOQNVFSCSVMHAEALHNHYTQKSLSLSPG 400
Db 205 VLSDSGSFFLYSKLTVDKSRWQOQNVFSCSVMHAEALHNHYTQKSLSLSPG 254
RESULT 4
S69339
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-2000
C;Accession: S69339
R;Hamlilchi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogné, M.
Eur. J. Biochem. 229, 54-60, 1995
A;Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A;Reference number: S69339; MUID:95262687
A;Accession: S69339
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-374 <RHA>
A;Cross-references: EMBL:X81695
C;Superfamily: immunoglobulin C region; immunoglobulin homology
Query Match 54.7%; Score: 1225; DB 2; Length 374;
Best Local Similarity 96.6%; Pred. No. 2.7e-68;
Matches 225; Conservative: 3; Mismatches 5; Indels 0; Gaps 0;
QY 168 SESTQKVDKTHTCPPCAPPELLGSPSVFLPPPKPDKTLMISRTPEVTCVVVDVSHEDPEV 227
Db 141 SSEPCKDTHTCPPCAPPELLGSPSVFLPPPKPDKTLMISRTPEVTCVVVDVSHEDPEV 200
QY 228 KFNWYDGVGVHNAKTPREEQNSTYRVSVLTVLHODWLNKKEYCKYKSNKALPAPIE 287
Db 201 KFNWYDGVGVHNAKTPREEQNSTYRVSVLTVLHODWLNKKEYCKYKSNKALPAPIE 260
QY 288 KTISKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 347
Db 261 KTISKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 320
QY 348 TTPVLDSDGSFFLYSKLTVDKSRWQOQNVFSCSVMHAEALHNHYTQKSLSLSPG 400
Db 321 TTPVLDSDGSFFLYSKLTVDKSRWQOQNVFSCSVMHAEALHNHYTQKSLSLSPG 373
RESULT 5
PT0207
Ig gamma chain C region - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C;Accession: PT0207
R;Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A;Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A;Reference number: PT0207; MUID:91287716
A;Accession: PT0207
A;Molecule type: mRNA
A;Residues: 1-234 <EHR>
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;48-117/Domain: immunoglobulin homology <IMM>
Query Match 52.8%; Score: 1183.5; DB 2; Length 234;
Best Local Similarity 94.4%; Pred. No. 5.7e-66;
Matches 221; Conservative: 1; Mismatches 3; Indels 9; Gaps 1;
QY 170 STQKVDK-----THTCPCAPPELLGSPSVFLPPPKPDKTLMISRTPEVTCVVVDV 220
Db 1 SNTKVDKRVKPKSCDTHTCPPCAPPELLGSPSVFLPPPKPDKTLMISRTPEVTCVVVDV 60
QY 221 SHEDPEVKFNWYDGVGVHNAKTPREEQNSTYRVSVLTVLHODWLNKKEYCKYKSNK 280

Db 61 SHEDPEVKFNWYDGVGVHNAKTPREEQNSTYRVSVLTVLHODWLNKKEYCKYKSNK 120
QY 281 ALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 340
Db 121 ALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 180
QY 341 PENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOQNVFSCSVMHAEALHNHYTQKS 394
Db 181 PENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOQNVFSCSVMHAEALHNHYTQKS 234
RESULT 6
A23511
Ig gamma-3 chain C region (allotype G3m(b)) - human
C:Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
C;Accession: A23511
R;Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
Nucleic Acids Res. 14, 1779-1789, 1986
A;Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene.
A;Reference number: A23511; MUID:86148507
A;Accession: A23511
A;Molecule type: DNA
A;Residues: 1-377 <HUC>
A;Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056
C;Genetics:
A;Gene: GDB:IGHG3
A;Cross-references: GDB:119339; OMIM:147120
A;Map position: 14q32.33-14q32.33
C;Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;20-85/Domain: immunoglobulin homology <IMM>
Query Match 51.5%; Score: 1154.5; DB 2; Length 377;
Best Local Similarity 67.1%; Pred. No. 5.7e-64;
Matches 228; Conservative: 19; Mismatches 32; Indels 61; Gaps 7;
QY 76 CNRTHNRVCKEGRYLEIEF-----CLKHRSC---PPFGVYVQAGTPERNYTC 121
Db 83 CNVNH-RPSNTKVDKRVELKTLGLDTHTCPCPEKSCDTPPP-----C 126
QY 122 KCPDGFESNETSSKAP-CRKHTNCSVFGLLLTQGNATHDNICSGNSESTQKVDKHTPC 180
Db 127 PRCEP-KSCDTPPCPCPEKSC-----DTDPCC 156
QY 181 PPCPAPELLGSPSVFLPPPKPDKTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVGVEN 240
Db 157 PRCPAPELLGSPSVFLPPPKPDKTLMISRTPEVTCVVVDVSHEDPEVQFKWYDGVGVEN 216
QY 241 AKTKPREEQNSTYRVSVLTVLHODWLNKKEYCKYKSNKALPAPIEKTISKAKGQPREP 300
Db 217 AKTKPREEQNSTYRVSVLTVLHODWLNKKEYCKYKSNKALPAPIEKTISKAKGQPREP 276
QY 301 QVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL 360
Db 277 QVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL 336
QY 361 YSKLTVDKSRWQOQNVFSCSVMHAEALHNHYTQKSLSLSPG 400
Db 337 YSKLTVDKSRWQOQNVFSCSVMHAEALHNHYTQKSLSLSPG 376
RESULT 7
A60764
Ig gamma-3 chain C region, form LAT - human
C:Species: Homo sapiens (man)
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999
C;Accession: A60764
R;Huck, S.; Lefranc, G.; Lefranc, M.P.
Immunogenetics 30, 250-257, 1989

A;Title: A human immunoglobulin IGHG3 allele (Gmb0, bl, c3, c5, u) with an IGHG4 convers

A;Reference number: A60764; MUID:90007613

A;Accession: A60764

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-377 <HUC>

C;Superfamily: immunoglobulin C region; immunoglobulin homology

C;Keywords: immunoglobulin

F;20-85/Domain: immunoglobulin homology <IMM>

Query Match 51.5%; Score 1152.5; DB 2; Length 377;
Best Local Similarity 67.1%; Pred.No.7.5e-64;
Matches 228; Conservative 19; Mismatches 32; Indels 61; Gaps 7;

QY 76 CNRTHNRVCECKEGRYLEIEF-----CLKHSC---PPFGVVOAGTPERTVC 121
||| :|: :|: :|: :|: ||
Db 83 CNNVNH-KPSNTKVDKRVELTPLGDTHTCPRCEPKSCDTPPP-----C 126
||| :|: :|: :|: :|:
QY 122 KRCPDGFFSNETSSKAP-CRKHTNCVSFGLLLLTQGNATHDNICSGNSESTQKVKDTHTC 180
||| :|: :|: :|: :|:
Db 127 PRCEPE-KSCDTPPCRCPEPKS-----DTPPCC 156
||| :|: :|: :|: :|:
QY 181 PPCPAPELLGGPSVFLEPPPKPDFLMTSRPEVTCVVVDVSHEDPEVKFNWYDGVVEHN 240
||| :|: :|: :|: :|:
Db 157 PRCPAPELLGGPSVFLEPPPKPDFLMTSRPEVTCVVVDVSHEDPEVQFKWYDGVVEHN 216
||| :|: :|: :|: :|:
QY 241 AKTKPREQINSYRVVSVTLVLHODWLNGKEYKCKYSNKALPAPIEKTISKAKGPREP 300
|||| :|: :|: :|: :|:
Db 217 AKTKPREQINSTRVSVTLVLHODWLNGKEYKCKYSNKALPAPIEKTISKTKGQPREP 276
|||| :|: :|: :|: :|:
QY 301 QVYTLPSRDRLTKNQVSLTCLVKGFYPSDIAEVESWGOPENNYKTTPTPVLDSDGSFFL 360
|||| :|: :|: :|: :|:
Db 277 QVYTLPSREEMTKNQVSLTCLVKGFYPSDIAEVESWGOPENNYNTPPVLDSDGSFFL 336
|||| :|: :|: :|: :|:
QY 361 YSKLTVDKSRWQQGNVPSCSMHEALHNHYTQKSLSLSPG 400
|||| :|: :|: :|: :|:
Db 337 YSRLTVDKSRWQEGNVFSCSMHEALHNRETQKSLSLSPG 376
|||| :|: :|: :|: :|:

RESULT 8

G2HU

Ig gamma-2 chain C region - human

C;Species: Homo sapiens (man)

C;Date: 30-Apr-1981 #sequence.revision 13-Jun-1983 #text_change 21-Jul-2000

C;Accession: A93906; A92809; A90752; A93132; A02148

R;Ellison, J.; Hood, L.

Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982

A;Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con

A;Reference number: A93906; MUID:82197621

A;Accession: A93906

A;Molecule type: DNA

A;Residues: 1-326 <ELL>

A;Cross-references: GB:J00230; GB:J00230; NID:g32759; PIDN:CAB58438.1; PID:g6066056

A;Note: Lys-326 is probably removed posttranslationally

R;Wang, A.C.; Tung, E.; Fudenberg, H.H.

J. Immunol. 125, 1048-1054, 1980

A;Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f

A;Reference number: A92809; MUID:81007873

A;Contents: myeloma protein Til

A;Accession: A92809

A;Molecule type: protein

A;Residues: 1-19,'Q','21-57,'Z','59','A','61-193,'D','195-325 <WAN>

A;Note: Trp-156 is at or near the complement-binding site

R;Connell, G.E.; Parr, D.M.; Hofmann, T.

Can. J. Biochem. 57, 758-767, 1979

A;Title: The amino acid sequences of the three heavy chain constant region domains of a

A;Reference number: A90752; MUID:80001357

A;Contents: myeloma protein Zie

A;Accession: A90752

A;Molecule type: protein

A;Residues: 1-24,'E','26-57,'EV','60-85;132-171,'zzz','175','B','177-193,'D','195-196,'Q','198-

R; Hofmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A; Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin G.
A; Reference number: A93132; MUID: 80114419
A; Contents: Zie
A; Accession: A93132
A; Molecule type: protein
A; Residues: 238-275 <HOF>
R; Hofmann, T.; Parr, D.M.
submitted to the Atlas. March 1980
A; Reference number: A94591
A; Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268
A; Note: the revised sequence differs from that shown in having 60-Ala and in the amid
ned
R; Milstein, C.; Frangione, B.
Biochem. J. 121, 217-225, 1971
A; Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.
A; Reference number: A90253; MUID: 72033500
A; Contents: annotation; myeloma protein Sa, disulfide bonds
R; Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A; Title: Structural studies of immunoglobulin G.
A; Reference number: A93157; MUID: 69064124
A; Contents: annotation; Sa, disulfide bonds
C; Genetics:
A; Gene: GDB: IGHG2
A; Cross-references: GDB: I19338; OMIM: 147110
A; Map position: 14q32.33-14q32.33
C; Complex: An immunoglobulin heterotetramer consists of two identical light (chain) disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into higher order structures by interchain (to heavy chain) #status experimental
C; Superfamily: immunoglobulin C region; immunoglobulin homology
C; Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin F; 20-85/Domain: immunoglobulin homology <IM1>
F; 133-202/Domain: immunoglobulin homology <IM2>
F; 239-306/Domain: immunoglobulin homology <IM3>
F; 14/Disulfide bonds: interchain (to light chain) #status experimental
F; 27-83, 140-200, 246-304/Disulfide bonds: #status experimental
F; 102, 103, 106, 109/Disulfide bonds: interchain (to heavy chain) #status experimental
F; 176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 51.3%; Score 1148.5; DB 1; Length 326;
Best Local Similarity 90.3%; Pred. No. 1, 1e-63;
Matches 214; Conservative 8; Mismatches 8; Indels 7; Gaps 2;

QY 170 STOKVKT-----HTCPPCPAPELGGPSVFLLPFPKPKDTLMISRTPEVTCVVVDVSHE 223
| | | | | : : : : :
Db 90 SNTKVKDVTVERKCCEPCPPAPP-VAGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHE 148

QY 224 DPEVKENWYDGVGEVINAKTKPREEOYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALP 283
| | | | | : : : : :
Db 149 DPEQFNWYDGVGEVINAKTKPREEFNFNSTRFRVSVLTIVVHODWLNGKEYCKVSNKGLP 208

QY 284 APIETKISKAGOPREPQQVYTLPPSRDELTKNOVSLTCLVKGFYPSPDIAVEWESNGOPEN 343
| | | | | : : : : :
Db 209 APIETKISKTKGPREFQVYTLPPSREEXTKNQVSLTCLVKGFYPSPDIAVEWESNGOPEN 268

QY 344 NYKTTTPVLDSGSFFLYSKLTVDKSRWQGNVFSQVMHEALHNHYTOKLSLSLSPG 400
| | | | | : : : : :
Db 269 NYKTTTPMLDSGSFFLYSKLTVDKSRWQGNVFSQVMHEALHNHYTOKLSLSLSPG 325

RESULT 9
G4HU
Ig gamma-4 chain C region - human
C; Species: Homo sapiens (man)
C; Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999
C; Accession: A90933; A02150
R; Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A; Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A; Reference number: A90933; MUID: 83157104
A; Accession: A90933

4

A;Molecule type: DNA
A;Residues: 1-327 <ELL>
A;Note: the sequence was determined from the germline gene
R;Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A;Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant
A;Reference number: A90249; MUID:70207560
A;Accession: A90249
A;Molecule type: protein
A;Residues: 1-30;81-326 <PIN>
C;Genetics:
A;Gene: GDB:IGHG4
A;Cross-references: GDB:119340; OMIM:147130
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 111/1; 221/1
A;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) and two heavy (lambda) chains. In some cases, such as IgA and IgM, the subunits associate into large aggregates. In other cases, such as IgG and IgE, the subunits remain as monomers.
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;99-110/Region: hinge
F;134-203/Domain: immunoglobulin homology <IM2>
F;240-307/Domain: immunoglobulin homology <IM3>
F;14/Disulfide bonds: Interchain (to light chain) #status experimental
F;27-83,141-201,247-305/Disulfide bonds: #status predicted
F;106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F;177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.7%; Score 1136; DB 1; Length 327;
Best Local Similarity 89.5%; Pred. No. 6.6e-63;
Matches 212; Conservative 8; Mismatches 11; Indels 6; Gaps 1;
QY 170 STQKVDKT-----HTCPCPAPPELLGGPSVFLPAPKDTLMISRTPEVTCVVVDVSH 223
Db 90 SNKYKRVESKYGPPCPAPPELLGGPSVFLPAPKDTLMISRTPEVTCVVVDVSH 149
QY 224 DPEYKFNWYDGVVEVHNNAKTPREEQNSTYRVSVLTSLVQLQDNLNGKPKCKVSNKALP 283
Db 150 DPEYKFNWYDGVVEVHNNAKTPREEQNSTYRVSVLTSLVQLQDNLNGKPKCKVSNKGLP 209
QY 284 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPEN 343
Db 210 SSIEKTIKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPEN 269
QY 344 NYKTPPVLDSDGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPG 400
Db 270 NYKTPPVLDSDGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSPG 326

RESULT 10
G3HUW1
Ig gamma-3 heavy chain disease proteins - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C;Accession: A90442; A92219; A90198; A93915; A02149
R;Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A;Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-chain disease protein
A;Reference number: A90442; MUID:81021548
A;Contents: heavy chain disease protein wis
A;Accession: A90442
A;Molecule type: protein
A;Residues: 1-289 <FRA>
A;Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain disulfide bond between residues 12 and 13 of the CH1 region. Residue 12 of the CH1 region is a cysteine which is not present in the reference sequence that follows
R;Michaelson, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A;Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication of the hinge region
A;Reference number: A92219; MUID:77118561
A;Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein wis
A;Accession: A92219

A;Molecule type: protein
A;Residues: 12-97 <MIC>
A;Note: the hinge region in gamma-3 chains is about four times as long as in other gamma-3 chains
R;Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A;Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of t
A;Reference number: A90198; MUID:77021516
A;Contents: heavy chain disease protein ZUC, partial sequence corresponding to residue
A;Accession: A90198
A;Molecule type: protein
A;Residues: 59-125, 'EB', 128-226, 228-289 <WOL>
A;Note: this protein lacks most of the V region, all of the CH1 region, and part of t
R;Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A;Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
A;Reference number: A93915; MUID:82247835
A;Contents: heavy chain disease protein Omim
A;Accession: A93915
A;Molecule type: mRNA
A;Residues: 12-70;72-114;116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157, 'E', 159-160, 'E', 162-163, 'E', 165-166, 'E', 168-169, 'E', 171-172, 'E', 174-175, 'E', 177-178, 'E', 180-181, 'E', 183-184, 'E', 186-187, 'E', 189-190, 'E', 192-193, 'E', 195-196, 'E', 198-199, 'E', 201-202, 'E', 204-205, 'E', 207-208, 'E', 210-211, 'E', 213-214, 'E', 216-217, 'E', 219-220, 'E', 222-223, 'E', 225-226, 'E', 228-229, 'E', 231-232, 'E', 234-235, 'E', 237-238, 'E', 240-241, 'E', 243-244, 'E', 246-247, 'E', 249-250, 'E', 252-253, 'E', 255-256, 'E', 258-259, 'E', 261-262, 'E', 264-265, 'E', 267-268, 'E', 270-271, 'E', 273-274, 'E', 276-277, 'E', 279-280, 'E', 282-283, 'E', 285-286, 'E', 288-289, 'E', 291-292, 'E', 294-295, 'E', 297-298, 'E', 300-301, 'E', 303-304, 'E', 306-307, 'E', 309-310, 'E', 312-313, 'E', 315-316, 'E', 318-319, 'E', 321-322, 'E', 324-325, 'E', 327-328, 'E', 330-331, 'E', 333-334, 'E', 336-337, 'E', 339-340, 'E', 342-343, 'E', 345-346, 'E', 348-349, 'E', 351-352, 'E', 354-355, 'E', 357-358, 'E', 360-361, 'E', 363-364, 'E', 366-367, 'E', 369-370, 'E', 372-373, 'E', 375-376, 'E', 378-379, 'E', 381-382, 'E', 384-385, 'E', 387-388, 'E', 390-391, 'E', 393-394, 'E', 396-397, 'E', 399-400, 'E', 402-403, 'E', 405-406, 'E', 408-409, 'E', 411-412, 'E', 414-415, 'E', 417-418, 'E', 420-421, 'E', 423-424, 'E', 426-427, 'E', 429-430, 'E', 432-433, 'E', 435-436, 'E', 438-439, 'E', 441-442, 'E', 444-445, 'E', 447-448, 'E', 450-451, 'E', 453-454, 'E', 456-457, 'E', 459-460, 'E', 462-463, 'E', 465-466, 'E', 468-469, 'E', 471-472, 'E', 474-475, 'E', 477-478, 'E', 480-481, 'E', 483-484, 'E', 486-487, 'E', 489-490, 'E', 492-493, 'E', 495-496, 'E', 498-499, 'E', 501-502, 'E', 504-505, 'E', 507-508, 'E', 510-511, 'E', 513-514, 'E', 516-517, 'E', 519-520, 'E', 522-523, 'E', 525-526, 'E', 528-529, 'E', 531-532, 'E', 534-535, 'E', 537-538, 'E', 540-541, 'E', 543-544, 'E', 546-547, 'E', 549-550, 'E', 552-553, 'E', 555-556, 'E', 558-559, 'E', 561-562, 'E', 564-565, 'E', 567-568, 'E', 570-571, 'E', 573-574, 'E', 576-577, 'E', 579-580, 'E', 582-583, 'E', 585-586, 'E', 588-589, 'E', 591-592, 'E', 594-595, 'E', 597-598, 'E', 600-601, 'E', 603-604, 'E', 606-607, 'E', 609-610, 'E', 612-613, 'E', 615-616, 'E', 618-619, 'E', 621-622, 'E', 624-625, 'E', 627-628, 'E', 630-631, 'E', 633-634, 'E', 636-637, 'E', 639-640, 'E', 642-643, 'E', 645-646, 'E', 648-649, 'E', 651-652, 'E', 654-655, 'E', 657-658, 'E', 660-661, 'E', 663-664, 'E', 666-667, 'E', 669-670, 'E', 672-673, 'E', 675-676, 'E', 678-679, 'E', 681-682, 'E', 684-685, 'E', 687-688, 'E', 690-691, 'E', 693-694, 'E', 696-697, 'E', 699-700, 'E', 702-703, 'E', 705-706, 'E', 708-709, 'E', 711-712, 'E', 714-715, 'E', 717-718, 'E', 720-721, 'E', 723-724, 'E', 726-727, 'E', 729-730, 'E', 732-733, 'E', 735-736, 'E', 738-739, 'E', 741-742, 'E', 744-745, 'E', 747-748, 'E', 750-751, 'E', 753-754, 'E', 756-757, 'E', 759-760, 'E', 762-763, 'E', 765-766, 'E', 768-769, 'E', 771-772, 'E', 774-775, 'E', 777-778, 'E', 780-781, 'E', 783-784, 'E', 786-787, 'E', 789-790, 'E', 792-793, 'E', 795-796, 'E', 798-799, 'E', 801-802, 'E', 804-805, 'E', 807-808, 'E', 810-811, 'E', 813-814, 'E', 816-817, 'E', 819-820, 'E', 822-823, 'E', 825-826, 'E', 828-829, 'E', 831-832, 'E', 834-835, 'E', 837-838, 'E', 840-841, 'E', 843-844, 'E', 846-847, 'E', 849-850, 'E', 852-853, 'E', 855-856, 'E', 858-859, 'E', 861-862, 'E', 864-865, 'E', 867-868, 'E', 870-871, 'E', 873-874, 'E', 876-877, 'E', 879-880, 'E', 882-883, 'E', 885-886, 'E', 888-889, 'E', 891-892, 'E', 894-895, 'E', 897-898, 'E', 900-901, 'E', 903-904, 'E', 906-907, 'E', 909-910, 'E', 912-913, 'E', 915-916, 'E', 918-919, 'E', 921-922, 'E', 924-925, 'E', 927-928, 'E', 930-931, 'E', 933-934, 'E', 936-937, 'E', 939-940, 'E', 942-943, 'E', 945-946, 'E', 948-949, 'E', 951-952, 'E', 954-955, 'E', 957-958, 'E', 960-961, 'E', 963-964, 'E', 966-967, 'E', 969-970, 'E', 972-973, 'E', 975-976, 'E', 978-979, 'E', 981-982, 'E', 984-985, 'E', 987-988, 'E', 990-991, 'E', 993-994, 'E', 996-997, 'E', 999-1000, 'E', 1002-1003, 'E', 1005-1006, 'E', 1008-1009, 'E', 1011-1012, 'E', 1013-1014, 'E', 1016-1017, 'E', 1019-1020, 'E', 1022-1023, 'E', 1025-1026, 'E', 1028-1029, 'E', 1031-1032, 'E', 1034-1035, 'E', 1037-1038, 'E', 1040-1041, 'E', 1043-1044, 'E', 1046-1047, 'E', 1049-1050, 'E', 1052-1053, 'E', 1055-1056, 'E', 1058-1059, 'E', 1061-1062, 'E', 1064-1065, 'E', 1067-1068, 'E', 1070-1071, 'E', 1073-1074, 'E', 1076-1077, 'E', 1079-1080, 'E', 1082-1083, 'E', 1085-1086, 'E', 1088-1089, 'E', 1091-1092, 'E', 1094-1095, 'E', 1097-1098, 'E', 1100-1101, 'E', 1103-1104, 'E', 1106-1107, 'E', 1109-1110, 'E', 1112-1113, 'E', 1115-1116, 'E', 1118-1119, 'E', 1121-1122, 'E', 1124-1125, 'E', 1127-1128, 'E', 1130-1131, 'E', 1133-1134, 'E', 1136-1137, 'E', 1139-1140, 'E', 1142-1143, 'E', 1145-1146, 'E', 1148-1149, 'E', 1151-1152, 'E', 1154-1155, 'E', 1157-1158, 'E', 1160-1161, 'E', 1163-1164, 'E', 1166-1167, 'E', 1169-1170, 'E', 1172-1173, 'E', 1175-1176, 'E', 1178-1179, 'E', 1181-1182, 'E', 1184-1185, 'E', 1187-1188, 'E', 1190-1191, 'E', 1193-1194, 'E', 1196-1197, 'E', 1199-1200, 'E', 1202-1203, 'E', 1205-1206, 'E', 1208-1209, 'E', 1211-1212, 'E', 1214-1215, 'E', 1217-1218, 'E', 1220-1221, 'E', 1223-1224, 'E', 1226-1227, 'E', 1229-1230, 'E', 1232-1233, 'E', 1235-1236, 'E', 1238-1239, 'E', 1241-1242, 'E', 1244-1245, 'E', 1247-1248, 'E', 1250-1251, 'E', 1253-1254, 'E', 1256-1257, 'E', 1259-1260, 'E', 1262-1263, 'E', 1265-1266, 'E', 1268-1269, 'E', 1271-1272, 'E', 1274-1275, 'E', 1277-1278, 'E', 1280-1281, 'E', 1283-1284, 'E', 1286-1287, 'E', 1289-1290, 'E', 1292-1293, 'E', 1295-1296, 'E', 1298-1299, 'E', 1301-1302, 'E', 1304-1305, 'E', 1307-1308, 'E', 1310-1311, 'E', 1313-1314, 'E', 1316-1317, 'E', 1319-1320, 'E', 1322-1323, 'E', 1325-1326, 'E', 1328-1329, 'E', 1331-1332, 'E', 1334-1335, 'E', 1337-1338, 'E', 1340-1341, 'E', 1343-1344, 'E', 1346-1347, 'E', 1349-1350, 'E', 1352-1353, 'E', 1355-1356, 'E', 1358-1359, 'E', 1361-1362, 'E', 1364-1365, 'E', 1367-1368, 'E', 1370-1371, 'E', 1373-1374, 'E', 1376-1377, 'E', 1379-1380, 'E', 1382-1383, 'E', 1385-1386, 'E', 1388-1389, 'E', 1391-1392, 'E', 1394-1395, 'E', 1397-1398, 'E', 1400-1401, 'E', 1403-1404, 'E', 1406-1407, 'E', 1409-1410, 'E', 1412-1413, 'E', 1415-1416, 'E', 1418-1419, 'E', 1421-1422, 'E', 1424-1425, 'E', 1427-1428, 'E', 1430-1431, 'E', 1433-1434, 'E', 1436-1437, 'E', 1439-1440, 'E', 1442-1443, 'E', 1445-1446, 'E', 1448-1449, 'E', 1451-1452, 'E', 1454-1455, 'E', 1457-1458, 'E', 1460-1461, 'E', 1463-1464, 'E', 1466-1467, 'E', 1469-1470, 'E', 1472-1473, 'E', 1475-1476, 'E', 1478-1479, 'E', 1481-1482, 'E', 1484-1485, 'E', 1487-1488, 'E', 1490-1491, 'E', 1493-1494, 'E', 1496-1497, 'E', 1499-1500, 'E', 1502-1503, 'E', 1505-1506, 'E', 1508-1509, 'E', 1511-1512, 'E', 1514-1515, 'E', 1517-1518, 'E', 1520-1521, 'E', 1523-1524, 'E', 1526-1527, 'E', 1529-1530, 'E', 1532-1533, 'E', 1535-1536, 'E', 1538-1539, 'E', 1541-1542, 'E', 1544-1545, 'E', 1547-1548, 'E', 1550-1551, 'E', 1553-1554, 'E', 1556-1557, 'E', 1559-1560, 'E', 1562-1563, 'E', 1565-1566, 'E', 1568-1569, 'E', 1571-1572, 'E', 1574-1575, 'E', 1577-1578, 'E', 1580-1581, 'E', 1583-1584, 'E', 1586-1587, 'E', 1589-1590, 'E', 1592-1593, 'E', 1595-1596, 'E', 1598-1599, 'E', 1601-1602, 'E', 1604-1605, 'E', 1607-1608, 'E', 1610-1611, 'E', 1613-1614, 'E', 1616-1617, 'E', 1619-1620, 'E', 1622-1623, 'E', 1625-1626, 'E', 1628-1629, 'E', 1631-1632, 'E', 1634-1635, 'E', 1637-1638, 'E', 1640-1641, 'E', 1643-1644, 'E', 1646-1647, 'E', 1649-1650, 'E', 1652-1653, 'E', 1655-1656, 'E', 1658-1659, 'E', 1661-1662, 'E', 1664-1665, 'E', 1667-1668, 'E', 1670-1671, 'E', 1673-1674, 'E', 1676-1677, 'E', 1679-1680, 'E', 1682-1683, 'E', 1685-1686, 'E', 1688-1689, 'E', 1691-1692, 'E', 1694-1695, 'E', 1697-1698, 'E', 1700-1701, 'E', 1703-1704, 'E', 1706-1707, 'E', 1709-1710, 'E', 1712-1713, 'E', 1715-1716, 'E', 1718-1719, 'E', 1721-1722, 'E', 1724-1725, 'E', 1727-1728, 'E', 1730-1731, 'E', 1733-1734, 'E', 1736-1737, 'E', 1739-1740, 'E', 1742-1743, 'E', 1745-1746, 'E', 1748-1749, 'E', 1751-1752, 'E', 1754-1755, 'E', 1757-1758, 'E', 1760-1761, 'E', 1763-1764, 'E', 1766-1767, 'E', 1769-1770, 'E', 1772-1773, 'E', 1775-1776, 'E', 1778-1779, 'E', 1781-1782, 'E', 1784-1785, 'E', 1787-1788, 'E', 1790-1791, 'E', 1793-1794, 'E', 1796-1797, 'E', 1799-1800, 'E', 1802-1803, 'E', 1805-1806, 'E', 1808-1809, 'E', 1811-1812, 'E', 1814-1815, 'E', 1817-1818, 'E', 1820-1821, 'E', 1823-1824, 'E', 1826-1827, 'E', 1829-1830, 'E', 1832-1833, 'E', 1835-1836, 'E', 1838-1839, 'E', 1841-1842, 'E', 1844-1845, 'E', 1847-1848, 'E', 1850-1851, 'E', 1853-1854, 'E', 1856-1857, 'E', 1859-1860, 'E', 1862-1863, 'E', 1865-1866, 'E', 1868-1869, 'E', 1871-1872, 'E', 1874-1875, 'E', 1877-1878, 'E', 1880-1881, 'E', 1883-1884, 'E', 1886-1887, 'E', 1889-1890, 'E', 1892-1893, 'E', 1895-1896, 'E', 1898-1899, 'E', 1901-1902, 'E', 1904-1905, 'E', 1907-1908, 'E', 1910-1911, 'E', 1913-1914, 'E', 1916-1917, 'E', 1919-1920, 'E', 1922-1923, 'E', 1925-1926, 'E', 1928-1929, 'E', 1931-1932, 'E', 1934-1935, 'E', 1937-1938, 'E', 1940-1941, 'E', 1943-1944, 'E', 1946-1947, 'E', 1949-1950, 'E', 1952-1953, 'E', 1955-1956, 'E', 1958-1959, 'E', 1961-1962, 'E', 1964-1965, 'E', 1967-1968, 'E', 1970-1971, 'E', 1973-1974, 'E', 1976-1977, 'E', 1979-1980, 'E', 1982-1983, 'E', 1985-1986, 'E', 1988-1989, 'E', 1991-1992, 'E', 1994-1995, 'E', 1997-1998, 'E', 1999-2000, 'E', 2002-2003, 'E', 2005-2006, 'E', 2008-2009, 'E', 2011-2012, 'E', 2014-2015, 'E', 2017-2018, 'E', 2020-2021, 'E', 2023-2024, 'E', 2026-2027, 'E', 2029-2030, 'E', 2032-2033, 'E', 2035-2036, 'E', 2038-2039, 'E', 2041-2042, 'E', 2044-2045, 'E', 2047-2048, 'E', 2050-2051, 'E', 2053-2054, 'E', 2056-2057, 'E', 2059-2060, 'E', 2062-2063, 'E', 2065-2066, 'E', 2068-2069, 'E', 2071-2072, 'E', 2074-2075, 'E', 2077-2078, 'E', 2080-2081, 'E', 2083-2084, 'E', 2086-2087, 'E', 2089-2090, 'E', 2092-2093, 'E', 2095-2096, 'E', 2098-2099, 'E', 2101-2102, 'E', 2104-2105, 'E', 2107-2108, 'E', 2110-2111, 'E', 2113-2114, 'E', 2116-2117, 'E', 2119-2120, 'E', 2122-2123, 'E', 2125-2126, 'E', 2128-2129, 'E', 2131-2132, 'E', 2134-2135, 'E', 2137-2138, 'E', 2140-2141, 'E', 2143-2144, 'E', 2146-2147, 'E', 2149-2150, 'E', 2152-2153, 'E', 2155-2156, 'E', 2158-2159, 'E', 2161-2162, 'E', 2164-2165, 'E', 2167-2168, 'E', 2170-2171, 'E', 2173-2174, 'E', 2176-2177, 'E', 2179-2180, 'E', 2182-2183, 'E', 2185-2186, 'E', 2188-2189, 'E', 2191-2192, 'E', 2194-2195, 'E', 2197-2198, 'E', 2200-2201, 'E', 2203-2204, 'E', 2206-2207, 'E', 2209-2210, 'E', 2212-2213, 'E', 2215-2216, 'E', 2218-2219, 'E', 2221-2222, 'E', 2224-2225, 'E', 2227-2228, 'E', 2230-2231, 'E', 2233-2234, 'E', 2236-2237, 'E', 2239-2240, 'E', 2242-2243, 'E', 2245-2246, 'E', 2248-2249, 'E', 2251-2252, 'E', 2254-2255, 'E', 2257-2258, 'E', 2260-2261, 'E', 2263-2264, 'E', 2266-2267, 'E', 2269-2270, 'E', 2272-2273, 'E', 2275-2276, 'E', 2278-2279, 'E', 2281-2282, 'E', 2284-2285, 'E', 2287-2288, 'E', 2290-2291, 'E', 2293-2294, 'E', 2296-2297, 'E', 2299-2300, 'E', 2302-2303, 'E', 2305-2306, 'E', 2308-2309, 'E', 2311-2312, 'E', 2314-2315, 'E', 2317-2318, 'E', 2320-2321, 'E', 2323-2324, 'E', 2326-2327, 'E', 2329-2330, 'E', 2332-2333, 'E', 2335-2336, 'E', 2338-2339, 'E', 2341-2342, 'E', 2344-2345, 'E', 2347-2348, 'E', 2350-2351, 'E', 2353-2354, 'E', 2356-2357, 'E', 2359-2360, 'E', 2362-2363, 'E', 2365-2366, 'E', 2368-2369, 'E', 2371-2372, 'E', 2374-2375, 'E', 2377-2378, 'E', 2380-2381, 'E', 2383-2384, 'E', 2386-2387, 'E', 2389-2390, 'E', 2392-2393, 'E', 2395-2396, 'E', 2398-2399, 'E', 2401-2402, 'E', 2404-2405, 'E', 2407-2408, 'E', 2410-2411, 'E', 2413-2414, 'E', 2416-2417, 'E', 2419-2420, 'E', 2422-2423, 'E', 2425-2426, 'E', 2428-2429, 'E', 2431-2432, 'E', 2434-2435, 'E', 2437-2438, 'E', 2440-2441, 'E', 2443-2444, 'E', 2446-2447, 'E', 2449-2450, 'E', 2452-2453, 'E', 2455-2456, 'E', 2458-24

RESULT 12
I47160

QY 125 PDGFFSNSTSSKAPC---RKHTNCSVFGLLLTQKGNAHDNICSGNSESTQKYDK---T 177
 | | | : | | | : | | | : | | | |
 Db 59 PGLSYLSMSMTVPASSLSKSYTCNV-----NHPATTTKYDKRVGT 100

Query Match	41.2%	Score	923	DB 1	Length	323
Best Local Similarity	61.5%	Pred.	No. 7.5e-50			
Matches	177	Conservative	33	Mismatches	46	Indels
QY	123	RCPDGPFSS----	NETSSKAPCRKHTNCSVFGLLITQKGNATHDNTCSGSEST			
Db	57	RQSSGLYSLSWSVSTSSQP----	VTGNV-----AHPATNT			
QY	178	--HTC--PPCAPPELLGGSPSVFLPPPKPDKTLMISRTPEVTCVVVDVSHEDPE				
Db	95	VAPTSCKPTCPPPELLGGSPSVFLPPPKPDKTLMISRTPEVTCVVVDVSDQDDPE				
QY	233	VDGVEVHNATKPREEQNSTYRVVSVLTVLHODWMLNGKEYCKKVSNKALPAPI				
		:				
Db	155	INNEQVTRAPPLREQFNSTIRVYSTLPIHODWMLRGKEPKCKVHNKALPAPI				
QY	293	AKGPREFQVYTLPPSRDELTKNQVSLTCLVKYGFPSPDIAVWESNGQPENNYK				
Db	215	ARGQPLEPKVYTMGPPEELSSRSVSLTCMINGFYPSDISVEWEKNGKAEDNYK				
QY	353	DSGSGFELYSKLTVDKSRWQGNQVFSCSMVHEALHNHYTQKSLSLSPG	400			
Db	275	DSGSGFELYNKLVSPTSWQRGDVFCTCSVMHEALHNHYTQKSLSRSPG	322			

Thu Mar 1 09:58:04 2001

us-09-389-782a-5.rpr

Page 8

Db 283 ADGSYFLXSLTVDKSAMWDQGTVTCTSVNHEALHNVHTOKAISRSPG 329

Search completed: March 1, 2001, 09:15:46
Job time: 144 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:00 ; Search time 70.93 Seconds
(without alignments)
192.831 Million cell updates/sec

Title: US-09-389-782A-5
Perfect score: 2240
Sequence: 1 ETFFPKYLHYDETSQLLC.....VMHEALHNHYTKSLSPG 400

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36.*
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1493	66.7	485	13	R24016
2	1488	66.4	518	15	R51003
3	1352	60.4	376	19	W60037
4	1337	59.7	376	18	W50287
5	1332	59.5	438	16	R81882
6	1332	59.5	438	19	W48976
7	1281	57.2	764	21	Y68949
8	1256	56.1	963	19	W70540
9	1252.5	55.9	388	19	W73513
10	1252.5	55.9	388	21	Y54063
11	1249.5	55.8	664	19	W71603
12	1249.5	55.8	664	21	Y80123

13	1249	55.8	445	20	Y24153
14	1248.5	55.7	664	19	W71604
15	1248.5	55.7	664	21	Y80124
16	1247	55.7	388	19	W73514
17	1247	55.7	388	21	Y54064
18	1245.5	55.6	482	19	W31646
19	1245.5	55.6	680	15	R48037
20	1243.5	55.5	449	20	W96278
21	1242	55.4	535	20	Y17414
22	1242	55.4	535	20	W70797
23	1242	55.4	532	21	Y92185
24	1241.5	55.4	375	18	W06683
25	1241	55.4	595	20	W86003
26	1240.5	55.4	387	17	R90920
27	1239	55.3	399	21	Y70867
28	1238.5	55.3	442	18	W10550
29	1238	55.3	396	18	W18574
30	1238	55.3	396	18	W18575
31	1237.5	55.2	784	21	Y92207
32	1237	55.2	347	13	R27163
33	1237	55.2	347	21	Y83136
34	1237	55.2	400	21	Y15123
35	1236	55.2	371	10	P93558
36	1236	55.2	460	21	Y69890
37	1236	55.2	480	20	W90206
38	1234.5	55.1	408	21	Y44461
39	1234	55.1	329	17	R91806
40	1234	55.1	351	14	R43685
41	1234	55.1	371	10	P91918
42	1234	55.1	446	17	W05829
43	1234	55.1	447	20	Y31669
44	1234	55.1	449	14	R43339
45	1234	55.1	449	19	W49816

ALIGNMENTS

RESULT 1	
R24016	
ID R24016 standard; Protein; 485 AA.	
XX R24016;	
XX AC	
XX DT 26-NOV-1992 (first entry)	
XX DE Fusion protein TNFRFC.	
XX KW Tumour necrosis factor; TNF; IgG1; immunoglobulin G1.	
XX OS Synthetic.	
XX FH Key	Location/Qualifier
FT Region	1..252
FT FT	/note= "human TNF extracellular receptor"
FT Region	253..270
FT FT	/note= "linker and hinge"
FT Region	271..379
FT FT	/note= "IgG1 CH2"
FT Region	380..485
FT FT	/note= "IgG1 CH3"
XX XX	
PN EP488170-A.	
XX XX	
PD 03-JUN-1992.	
XX XX	
PF OX40/Fc mut. C	
XX XX	
PR Integrin beta-1 ch	
XX XX	
PA Amino acid sequenc	
XX XX	
PI Laufer L, Oquendo P, Zettlmeissl G;	


```

|||||
Db 368 nstyrsvsvltvlgdwlngdyckvsnkalpampqktiskakgqprepyvtlppsrd 427
QY 311 ELTKNOVSLTCLVKGYPSPDIWESNGOPENNYKTTTPVLDSDGSFELYSKLTVDKSR 370
Db 428 eltkngvslclvkgfyrhriawesngppennnyktptpvidsdgsfflyskltvdksr 487
QY 371 WQGNVFCSCVMHEALHNYTKSLSPG 400
Db 488 wqgnvfscvmhealhnhytkslspg 517

RESULT 3
ID W60037 standard; Protein; 376 AA.
XX
AC W60037;
XX
DT 11-SEP-1998 (first entry)
XX
DE Antigenic peptide hFas (nd29) containing Fc region.
XX
KW Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
KW heart failure; kidney failure; graft-versus-host disease; antibody;
KW myocardial infarction; ischemic restenosis; endotoxemic shock.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT Protein
FT /note= "hFas (nd29) protein"
XX
PN W09818487-A1.
XX
PD 07-MAY-1998.
XX
PF 31-OCT-1997; 97WO-JP03978.
XX
PR 26-SEP-1997; 97JP-0262521.
PR 31-OCT-1996; 96JP-0290459.
PR 27-DEC-1996; 96JP-0351718.
XX
PA (MOCH ) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
PI Nagata S, Suda T, Yatomi T;
XX
DR WPI; 1998-271925/24.
DR N-PSDB; V34430.
XX
PT Use of Fas antagonist for treatment and prevention of
PT apoptosis-related diseases - such as heart or kidney failure,
PT graft-versus-host disease or liver disease
XX
PS Examples; Fig 5-9; 86pp; Japanese.
XX
CC This represents the antigenic peptide hFas (nd29) containing the Fc
CC region. The invention provides the use of Fas antagonist as an agent for
CC the treatment and prevention of apoptosis-related diseases. The Fas
CC antagonist can be a partial Fas antigen peptide containing the
CC extracellular part of the protein, but lacking the signal sequence, an
CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is
CC preferably a humanised antibody. The Fas antagonist is used in the
CC treatment and prevention of diseases such as myocardial infarction, heart
CC failure, ischemic heart disease, acute kidney failure, graft-versus-host
CC disease, ischemic restenosis of the heart, liver or kidney, and
CC endotoxemic shock, and also as an organ preservative in transplantation.
CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand
XX system.
XX Sequence 376 AA;
XX

```

```

Query Match 60.4%; Score 1352; DB 19; Length 376;
Best Local Similarity 67.3%; Pred No. 5e-78;
Matches 270; Conservative 15; Mismatches 60; Indels 56; Gaps 9;

QY 8 LHYDEFTSHQLLDCPKPPGTYLKQCHCTAKW-KTVCAPCPD-HYYTDSWHTSDCLYCSVP 65
Db 23 lhhdgqfchk---pcppgkerkardctvngdepdcvqdegkeytdkafsskcr-cr-l 77
QY 66 CKELQYVVKQF--CNRTHNRVCECKEGRYLEIEFCLKHRSCPP-----GFGVVQAGTPERNT 119
Db 78 cdeghgleveinctrtgntkcrckpnfcnstvc---ehcdpctkcehglikeitstnt 134
QY 120 VCKRCPCDGFNETSSKAPCRKHTNCSVFGLLLTLQKGNATHDNCNSENSESTQKVOKTHT 179
Db 135 kcke--egrsnepks-----cdktht 154
QY 180 CPPCPAPELIGGFSVFLFPKPKDMLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVVH 239
Db 155 cpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdvshdpevkfnwyvdgveh 214
QY 240 NAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISAKGQPRE 299
Db 215 naktkpreegynstyrsvsvltvlgdwlngkeykckvsnkalpapielktiskakgpre 274
QY 300 PQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIWESNGOPENNYKTTTPVLDSDGSFF 359
Db 275 pqvyltppsrdeltkngvslclvkgfypsdiawesngppennnyktptpvidsdgsff 334
QY 360 LYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSPG 400
Db 335 lyskltvdksrwqgnvfscvmhealhnhytkslspg 375

RESULT 4
W50287
ID W50287 standard; Protein; 376 AA.
XX
AC W50287;
XX
DT 16-JUL-1998 (first entry)
XX
DE Human Fas antigen derivative/IgG1 Fc fusion.
XX
KW Human; Fas antigen; derivative; apoptosis regulation; gene therapy;
KW treatment; diabetes; arthritis; lupus; hepatitis; Influenza; HIV;
KW apoptosis modulation; immunoglobulin G1 Fc; IgG1 Fc; fusion.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT /label= sig_peptide
FT /label= mat_peptide
XX
PN W09742319-A1.
XX
PD 13-NOV-1997.
XX
PF 01-MAY-1997; 97WO-JP01502.
XX
PR 02-MAY-1996; 96JP-0135760.
XX
PA (MOCH ) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
XX
PI Nagata S, Nakamura N;
XX
DR WPI; 1997-558981/51.
DR N-PSDB; V07004.
XX

```

PT Fas antigen derivative containing modified extracellular region -
PT has low antigenicity, promotes apoptosis and is useful in treatment
XX of viral and other diseases

PS Disclosure; Fig 4; 102pp; Japanese.

XX The present sequence is a Fas antigen derivative/IgG1 Fc
CC fusion, which contains a Fas antigen extracellular region lacking
CC one or more amino acid residues in the region from the
CC amino-terminal to (but excluding) the 1st cysteine residue
CC (preferably at least 29 residues are deleted).

CC The derivative is an effective regulator of apoptosis and can be
CC used (either by administration of the polypeptide, or by the use
CC of the coding DNA in gene therapy) to treat a range of diseases,
CC e.g. diabetes, arthritis, lupus and in particular viral diseases
CC such as hepatitis, influenza and HIV, by modulating apoptosis of
CC virus-infected cells.

XX Sequence 376 AA;

Query Match 59.78; Score 1337; DB 18; Length 376;
Best Local Similarity 66.68; Pred. No. 4.8e-77;
Matches 267; Conservative 16; Mismatches 62; Indels 56; Gaps 9;

QY 8 LHYDEETSHOLLCDKCPGTYLKQHTAKW-KTVACAPCPD-HYVTDGSHWTSDECLYCSVP 65

DB 23 lhhdgqfchk---pppggarkardctvngdepdcvpcqeqkeytdkahfsskrrcr-i 77

QY 66 CKELQYVQKE--CNRTNHRVCECKEGRYLETEFLKHKRSCPP---GFGVVQAGTSPRNT 119

DB 78 cdegheieveincrtntqtkrcpknfcnstvc---ehcdpctkcehgliiecktltsnt 134

QY 120 VCKRCPDGFSNETHSSKAPCKRHTNCVFGLLLTQKGNATHDNCNSBSTQKVDTHT 179

DB 135 kcke--egrsneks-----cdktht 154

QY 180 CPPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 239

DB 155 cpcpapellggpsvflfpkpkdtlmsrtpevtcvvdvshedpevkfnwvydgvvvh 214

QY 240 NAKTKPREEQYNSTYRVVSVLTCLVKGFPSPDIAVEWESNGPENNKTTPVLDSDGSEF 299

DB 215 naktkpreeqynstyrsvsvltvlgqdwlngkeykckvsnkalpapietkiskagqpre 274

QY 300 POVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEWESNGPENNKTTPVLDSDGSEF 359

DB 275 pqvylppsrdeltknqvsltcvlgkfyfypsdiavewesngqpennkttpvldsdgssff 334

QY 360 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSPG 400

DB 335 lysnltvdksrwqgnvfscsvmhealhnhytqkslsisp 375

RESULT 5

R81882

ID R81882 standard; Protein; 438 AA.

XX AC R81882;

XX 30-MAR-1996 (first entry)

XX Plasmid pDC406/OX40/Fc* encoding an OX40/Fc mutein protein.

XX OX40: OX40-L; cytokine; cell surface molecule; plasmid;

XX pDC406/OX40/Fc*; membrane glycoprotein.

XX Synthetic.

XX OS US5457035-A.

XX PN 10-OCT-1995.

XX PD

XX

PF 23-JUL-1993; 93US-0097827.

XX PR 23-JUL-1993; 93US-0097827.

XX (IMMV) IMMUNEX CORP.

XX Baum PR, Fanslow WC, Gayle RB, Goodwin RG;

XX WPI; 1995-357992/46.

DR N-PSDB; T00829.

XX New isolated DNA encoding the OX40 ligand polypeptide - also vectors

PT and host cells, used to produce recombinant ligand used in e.g.

PT prim. T cell culture, to modulate immune response etc.

XX Example 2; Column 35-38; 26pp; English.

XX This plasmid encodes an OX40/Fc antibody fragment mutein protein,

CC and is used to express a soluble OX40/Fc mutein fusion protein for

CC use in detecting cDNA clones encoding a OX40 ligand. The Fc

CC fragment may be derived from human IgG1, and the plasmid may be

CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell

CC line. Culture supernatant was purified by affinity chromatography

CC and this was used, together with labeled goat anti-human IgG to

CC screen various cell lines.

XX Sequence 438 AA;

QY Query Match 59.5%; Score 1332; DB 16; Length 438;

DB Best Local Similarity 62.4%; Pred. No. 1.2e-76;

Matches 262; Conservative 23; Mismatches 75; Indels 60; Gaps 6;

QY 20 CDKCPPTTYLKQHTAKWKTVCAPCPDPHYTD--SWHTSDECLYCSVCKELQVVKQECN 77

DB 39 cecqpgdgmvrchtrdtlchpccetgfyneavnytdckqctqcnh--rsgsalkqnc 96

QY 78 RTHNRVCECKEGRYLETEFLKHKRSCPPGFGVQAGTPERNVCKRCPCDGFSSNETSSA 137

DB 97 ptqdtvcr-----pgtqrdsgykgldcvpcppghfs--pgnng 137

QY 138 PCRKHTNCVFGLLLTQKGNATHDNCNSGNS-----EST----- 171

DB 138 acpwtntctlsqktrhpsasdlavcdsrllatlletqrptfrpttvgstvwprts 197

QY 172 -----QKVDKHTTCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDV 220

DB 198 elpstptliveprscdkthtccpccapeagapsvflfpkpkdtlmsrtpevtcvvvdv 257

QY 221 SHEDPEVKFNWYVDGVEVHNAKTTPREEQYNSTYRVVSVLTCLVKGFPSPDIAVEWESNGQ 280

DB 258 shedpevkfnwvydgvvvhnaaktppreeqynstyrsvsvltvlgqdwlngkeykckvsnk 317

QY 281 ALPAPIEKTISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEWESNGQ 340

DB 318 alpapietkiskagqpreqvytlppsrdeltknqvsltcvlgkfyfypsdiavewesngq 377

QY 341 PENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSPG 400

DB 378 pennykttppvldsdgssfflyskltvdksrwqgnvfscsvmhealhnhytqkslsisp 437

RESULT 6

W48976

ID W48976 standard; Protein; 438 AA.

XX AC W48976;

XX 25-SEP-1998 (first entry)

XX OX40/Fc mutein.

XX OX40/Fc; cytokine; T cell antigen; TH-2 immune response; OX40-L;

XX KW

chimeric.
 Chimeric - Homo sapiens.
 Chimeric - Mus sp.

Key Location/Qualifiers
 Region 1..206
 Region 207..438
 Note= "Extracellular domain of mouse OX40"
 Note= "Mutant Fc region of human IgG1 antibody"
 Misc-difference 225
 Note= "changed from Leu in wild-type to Ala in mutant"
 Misc-difference 226
 Note= "changed from Leu in wild-type to Gly in mutant"
 Misc-difference 228
 Note= "changed from Gly in wild-type to Ala in mutant"

US5783665-A.
 21-JUL-1998.
 22-JUN-1995; 95US-0494574.
 23-JUL-1993; 93US-0097827.
 22-JUN-1995; 95US-0494574.
 (IMMUNEX CORP.)

Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
 WPI: 1998-427099/36.
 N-PSDB; V32636.

Purified polypeptide OX-40 ligands - for co-stimulation of T-cell production and binding assays for OX-40 and homologues
 Example 2; Col 37-40; 26pp; English.

The present sequence represents the OX40/Fc fusion protein that contains the extracellular domain of mouse OX40 fused to the mutated Fc region of the human IgG1 antibody. The fusion protein was used for detecting cDNA clones encoding an OX40 ligand. The invention claims for a murine OX40-L cytokine (W48975) that binds to the murine T cell antigen, OX40. The OX40-L protein is claimed to be useful for co-stimulation of T-cell production and in binding assays for detecting OX40 or its homologues. The OX40-L protein is also claimed to generate a TH-2 immune response.

Sequence 438 AA;

Query Match 59.5%; Score 1332; DB 19; Length 438;
 Best Local Similarity 62.4%; Pred. No. 1.2e-76;
 Matches 262; Conservative 23; Mismatches 75; Indels 60; Gaps 6;

20 CDKCPGTYLKQCTAKWTCVAPCPDPHYTD--SWHTSDCLYCSFVCKELQYVKQECN 77
 39 cecqpggmvrndhtrdtlchpcetgfyneavnydtkcqtqcnh--rsgselkqnc 96
 78 RTHNRVCECKEGRYLEFCLKHSRCPGCVVQAGTPERTVCKRCPDGFNFSSKA 137
 97 ptdqtvrcr-----pgtqrqdsqgylgvdcpvcpqghfs--pqnq 137
 138 PCRKHTNCSVFGLLAQKGNATHNDCSGNS-----EST----- 171
 138 ackpwnctslgkqrhpsasdlavcedrslatlwtqrtrftrttvqsttvprrts 197
 172 -----QKVDKTHTCPPCAPPELLGGPSVFLFPKPKDLMISRTPEVTCVVDV 220
 198 elpstpvlveprscdkthtccpcaeaegapsvfifpkpkdmlmistrptevtcvvdv 257

221 SHEDPEVKFNWYDGVGVHNAKTKPREOYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 280
 258 shedpevkfnwYDGVGVHNAKTKPREOYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 317
 281 ALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQ 340
 318 alpapiEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQ 377
 341 PENNYKTPPVLDSGDSFFLYSKLTVDKSRWQQGNVPCSCVMHEALHNHYTQKSLSLSPG 400
 378 pennYKTPPVLDSGDSFFLYSKLTVDKSRWQQGNVPCSCVMHEALHNHYTQKSLSLSPG 437

RESULT 7
 Y68949
 ID Y68949 standard; Protein; 764 AA.
 XX
 AC Y68949;
 DT
 DE
 XX 30-MAY-2000 (first entry)
 DE Fusion protein of murine delta-related protein and human IgG Fc.
 XX
 KW Cell development cycle; Delta family; membrane surface-bound ligand;
 KW endothelial cell biology; gene therapy; subcortical infarct;
 KW cerebral autosomal dominant arteriopathy; leucoencephalopathy;
 KW ischemic stroke; chimera.
 XX
 OS Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 FH Key Location/Qualifiers
 FT Protein. 1..529
 FT /note= "extracellular region of the murine
 FT delta-related protein"
 FT Protein 533..764
 FT /note= "human IgG Fc portion"
 XX
 PN WC200006726-A2.
 XX
 PD 10-FEB-2000.
 XX
 PF 12-JUL-1999; 99WO-US15710.
 XX
 PR 27-JUL-1998; 98US-0123168.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Shutter JR, Stark KL;
 XX
 DR WPI: 2000-195294/17.
 DR N-PSDB; Z60926.
 XX
 XX Cell development cycle protein of delta family useful for treating
 PT various disorders associated with central nervous system e.g. cerebral
 PT autosomal dominant arteriopathy and ischemic strokes
 XX
 PS Example 6; Page 169-171; 171pp; English.
 XX
 CC The present sequence represents a fusion protein of the extracellular
 CC domain of a murine polypeptide, which is a member of the cell
 CC development cycle protein family known as the Delta family of
 CC mammalian membrane surface-bound ligands, and the human immunoglobulin
 CC (IgG) Fc portion. The murine delta-related protein gene is expressed
 CC within vascular endothelium indicates a role for the polypeptides
 CC in the control of endothelial cell biology. The murine polynucleotide
 CC is identified from a white adipose tissue cDNA library. The polypeptide
 CC is useful for identifying receptors, which bind to and/or are activated
 CC by the polypeptide. The polynucleotide is useful in gene therapy of
 CC cerebral autosomal dominant arteriopathy with subcortical infarcts and
 CC leucoencephalopathy, an autosomal dominant disorder causing ischemic
 CC strokes.

```

xx SQ Sequence 764 AA;
Query Match 57.28; Score 1281; DB 21; Length 764;
Best Local Similarity 59.6%; Pred. No. 3.6e-73;
Matches 270; Conservative 23; Mismatches 66; Indels 94; Gaps 15;
QY 11 DEETSHOLLCDKCPGPGYVYLKQCHTAKWKTCA-----PCPDHY 49
Db 342 dqensyhcl---cpqg-yvgqchshlt-cadspcfnggscrnagssyacecpnft 396
QY 50 TDSWH-----TSDECLYCSVPCKELQYVKQECNRTHNRVCECKEG-----RYLEIEFCL 98
Db 397 gsncekkvdrtcsnpgang-----qclnrgpsrtcrpgftgtthcelhsdca 446
QY 99 KHRSCPPGFGVVGAGT---PERNTVCKRCPDGF-----FSNETSKAPCRKHTNCSV 147
Db 447 r-spcang-----gtchdlengpvc-tcpagfsgrrcevrithdacasgpcfnagtc-- 496
QY 148 FGLLLTQGNATHDNICS-----GN-----SESTQKVDKTHTCPPCPAPE 187
Db 497 -----ytlgspnfvncpvygvsrcefpvgllpsfpaaepkscdkthtcpcpape 550
QY 188 LLGSPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNNAKTPRE 247
Db 551 llgspsvflppkpkdtlmsrtpevtcvcvvdvshknpvfnfnyvgvvevhnaktkpre 610
QY 248 EGYNSTRVSVLVHLQDWLNGEYKCKVSNKALPAPIETISKAKGQPREPOVYVTLPP 307
Db 611 eqynstyrsvsvltvlhqdwlngkyckvsnkalpapiektiskakgqprepqvtytlpp 670
QY 308 SRDELTKNOVSLTCLVKGFPYSDIAVEWESNGQPENNYKTPPVLDSDGSPFLYSLKLTVD 367
Db 671 srdeltnqsvltclvkgfypsdiavevesngqpennyyktppvldsdgspfllyskltvd 730
QY 368 KSRWQQGNVFCFSVMHEALHNYHTQKLSLSPG 400
Db 731 ksrwqqgnvfscsvmhealhnhytqkslsispg 763

RESULT 8
W0540
ID W0540 standard; Protein; 963 AA.
AC W0540;
XX W0540;
XX W0540;
DT 26-JAN-1999 (first entry)
XX Integrin beta-1 chain.
XX Integrin; beta-1 chain; immunoglobulin; chimeric; heterodimer complex;
XX inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent;
XX human.
XX Homo sapiens.
XX OS
XX FH Location/Qualifiers
XX Key 1..20
XX Peptide /note= "signal peptide"
XX Protein 21..963
XX /note= "mature protein"
XX PN W09832771-A1.
XX PD 30-JUL-1998.
XX PF 29-JAN-1998; 98W0-JP00370.
XX PR 29-AUG-1997; 97JP-0234544.
XX PR 29-JAN-1997; 97JP-0015118.
XX (TORA ) TORAY IND INC.

```

Kainoh M, Tanaka T;
WPI: 1998-427881/36.
N-PSDB; V33773.

Integrin-immunoglobulin chimeric protein heterodimer complexes as platelet substitutes - contain the alpha and beta integrin chains associated in stable state and bind to extracellular matrix in the presence of plasma components

Claim 9; Pages 50-57; 87pp; Japanese.

This represents an integrin beta-1 chain. The invention provides integrin-immunoglobulin chimeric protein heterodimer complexes that comprise an integrin alpha or beta chain associated with an immunoglobulin light or heavy chain. These chimeric proteins form heterodimer complexes, in particular with a chimeric protein containing an integrin alpha chain and an immunoglobulin chain with a chimeric protein containing an integrin beta chain and an immunoglobulin chain; the immunoglobulin chain in each case may be a heavy chain, or one of the two may be a light chain. The integrin alpha chain is preferably alpha 4 or alpha 2 and the integrin beta chain is preferably beta 1. Animal cells transformed with vectors containing the DNA coding for the above chimeric proteins can be used in the preparation of the chimeric proteins and their heterodimer complexes. The heterodimer complexes, which are useful for testing potential promoters and inhibitors of the binding of integrins to their ligands, function as blood platelet substitutes and hemostatics and as diagnostic agents.

xx SQ Sequence 963 AA;

Query Match 56.1%; Score 1256; DB 19; Length 963;
Best Local Similarity 51.4%; Pred. No. 1.7e-71;
Matches 272; Conservative 31; Mismatches 90; Indels 136; Gaps 14;

QY 1 ETFPKYLHYDEETS--HQLLCD-----KCPPTG-----YLRQHC 33
Db 441 dsfkiplgtfevevilqyceceqseqipekchegntfecagacnegrgrhc 500
QY 34 TAKWKTVCAPCPDHY-----YTDSWHTSDECLYCSVPCKELQYVKQ-----EC 76
Db 501 ecstdevnsedmdaycrkensseicnngcevcgqcrkrdrntneilysgkfcednfn 560
QY 77 NRTN-----RVCECKEGRYLEIEFC-LKHSCPP-----GGVQAGTPER 117
Db 561 drsngllcgngvckrcvcecpnytgscadcsldtceasngqicnrgicecg---- 616
QY 118 NTVCKRCPDGFFSNETSSKAP-----CRKH-----TNCVFGLLLT 153
Db 617 --vck-ctdpkfgqgtcemqctclgvcaehkecvqcrfankgekkdtctqcsyfnitkv 673
QY 154 QKGNATHDNI-----CSGNS-----EST 171
Db 674 esrdklpqvqpdpvshckekdvddcwfyfytysvngnmvnmhvvenpecptgpedpeep 733
QY 172 QKVDKTHTCPCPAPPELLGGPSVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFNW 231
Db 734 kscdkthtccpccpapelggpsvflfpkpkdtlmsrtpevtcvcvvdvshdpevkfnw 793
QY 232 YVDCVEVHNNAKTPREQYNSYTRVSVLVHLQDWLNGEYKCKVSNKALPAPIETIS 291
Db 794 yvdgvevhnaktpreeqynstyrsvsvltvlhqdwlngkyckvsnkalpapietis 853
QY 292 KAKGQPREPOVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 351
Db 854 kagqprepqvtytlppsrdeltnqvsitclvkgfypsdiavevesngqpennyyktppv 913
QY 352 LDSGSEFLYSLKLTVDKSRWQQGNVFCFSVMHEALHNYHTQKLSLSPG 400
Db 914 ldsdgsfilyskltvdksrwqqgnvfscsvmhealhnhytqkslsispg 962

CC glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;
 CC collagen vascular disease, e.g. systemic sclerosis, polymyositis,
 CC scleroderma, dermatomyositis or systemic lupus erythematosus; and
 CC fibrosis associated with restenosis. It is also used for treating
 CC wounds, to prevent overproduction of connective tissue and so prevent
 CC adhesions or scarring, and to prevent post-radiation fibrosis (by
 CC administration to patients about to undergo radiation therapy).
 XX
 SQ Sequence 388 AA;

Query Match 55.9%; Score 1252.5; DB 21; Length 388;
 Best Local Similarity 65.7%; Pred. No. 1e-71;
 Matches 257; Conservative 17; Mismatches 70; Indels 47; Gaps 10;

QY 37 WKTVCAPCPDHY---TDSWHTSDE-----CLYC---SPVCKELOYVVKQECNRTHN 81
 Db 17 wtrlastipphvqsvnmndmvtddngavkfpqickfcdvrsctcdnqkscmscsit-- 74
 QY 82 RVCCKEGRYLEIEFCLKHRSCPPGFGVVQAGTPERNVCKRCPPD-----GFFSNETS 135
 Db 75 sice-----kahevcv---avwrkndenitletvchdpklayhgfiledsas 118
 QY 136 KAPCRKHTNCSVFG---LLLKQGNATHDNTCSGSESTQK---VDKTHTCPPCPAPELL 189
 Db 119 pkcimkek--kvgetffmcscstdecdndhiifseeeyttspdlvdkthtccppapeall 176
 QY 190 GGPSVFLPPKPKDTLMSRPETVTCVVVYSHEDPEVKFNWYDGVVHNAKTPREEQ 249
 Db 177 ggsavilppkpkdtlmsrtpevtcvvvdshedpevkfnwvdygvevhnaktpreeq 236
 QY 250 YNSYRVVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLPPSR 309
 Db 237 ynstyrsvsvltvlhqdwlngkeyckvsnkalpapiektiskakgqpreqvylppsr 296
 QY 310 DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKS 369
 Db 297 deltknqvslclvkgyfypsdiawesngqpennnykttppvldsdgsfflyskltvdk 356
 QY 370 RWQGNVFCSCVMHEALHNHYTQKSLSLSPG 400
 Db 357 rwqgnvfscsvmhealhnhytqkslsispg 387

RESULT 11
 W1603
 ID W1603 standard; Protein; 564 AA.
 AC W1603;
 XX
 XX 19-NOV-1998 (first entry)
 DT Human neuturin receptor alpha/Fc sequence (If2a) fusion protein.
 DE
 XX Human; neuturin receptor alpha; NTNR-alpha; variant; chimeric;
 KW fusion protein; immunoadhesion; ret-expressing cell; neurological;
 KW renal; haematological disease.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN W09836072-A1.
 XX
 PD 20-AUG-1998.
 XX
 PF 17-FEB-1998; 98WO-US03179.
 XX
 PR 24-OCT-1997; 97US-0957063.
 PR 18-FEB-1997; 97US-0802805.
 PR 09-JUN-1997; 97US-0871913.
 XX
 PA (GETH) GENENTECH INC.
 XX

PI Hynes MA, Klein RD, Rosenthal A;
 DR WPI: 1998-467175/40.
 DR N-PSDB; V58006.
 XX
 PT New polypeptide(s) based on human neuturin receptor alpha and
 PT related nucleic acid - useful for increasing survival of
 PT ret-expressing cells for treating e.g. neurological, renal and
 PT haematological diseases
 XX
 PS Example 3; Page 81-83; 116pp; English.
 XX
 CC The present sequence represents human neuturin receptor alpha
 CC (NTNR-alpha)/Fc sequence (If2a) fusion protein, from an example of the
 CC present invention. NTNR-alpha proteins can be used: (a) to identify
 CC molecules that bind specifically to it (potential agonists and
 CC antagonists) and to purify such compounds; (b) to modulate response of
 CC cells to neuturin (NTN); (c) to increase survival of Ret-expressing
 CC cells or to activate Ret on the surface of cells (soluble glial derived
 CC neurotrophic factor receptor (GDNFR alpha) may be used the same way);
 CC (d) to increase the half-life of cognate ligands (especially NTN); (e)
 CC diagnostically to determine serum levels of its ligands; and (f) as
 CC animal feed additive or molecular weight marker. NTNR-alpha, its genes,
 CC (ant)agonists and antisense nucleic acids, are useful in vivo or ex vivo
 CC for treating conditions related to abnormal NTN activity or response,
 CC particularly neurological (central or enteric), renal or haematopoietic
 CC (spleen) diseases or injuries. Ab may be agonists or antagonists for
 CC therapeutic use (e.g. as antagonists to treat excessive/unwanted NTNR-
 CC alpha expression, e.g. in some tumours), also reagents for immunoassay
 CC and affinity purification.
 XX
 SQ Sequence 664 AA;

Query Match 55.8%; Score 1249.5; DB 19; Length 664;
 Best Local Similarity 63.6%; Pred. No. 3e-71;
 Matches 264; Conservative 23; Mismatches 57; Indels 71; Gaps 14;

QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDECLYCSPVCKELQYV 72
 Db 274 ncrasyqlvts-cpadnyqaclysyagmigfdmtpnyvdsptg---ivspwc----- 323
 QY 73 KOECNRTHNRVCECKEGRYLEIEF---CLKH-----RSCP--PGFGVVQAGTP 115
 Db 324 --scrgsgnmeece--kflr-dftenpclrnaiaqafngtdvnnvsgkpsfqatqprv 378
 QY 116 ERNVTVCRCPDGFFSNETSSKAPCRKHTNCSVFGLLLTQGNATHD-----NICS 165
 Db 379 ekt---pslpddlsdstslgtsttctsvqeggl----kannskelsmcfelttnlip 431
 QY 166 GNSESTQKVDKTHTCPPCPAPELLGSPVFLFPKPKDTLMSRTPEVTCVVVDVSHEDP 225
 Db 432 gprd---pvdkthccppcpapellggpsvflfppkpkdtlmsrtpevtcvvvdshedp 488
 QY 226 EVKENWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVTLHQDLNGKEYCKVSNKALPAP 285
 Db 489 evkfnwyvdgvevhnaktkpreeqynstyrsvsvltvlhqdwlngkeyckvsnkalpap 548
 QY 286 IEKTYISKAKGQPREPQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 345
 Db 549 iektiskakgqpreqvtytlppsreemtknqvsitclvkgyfypsdiawesngqpenny 608
 QY 346 KTTTPVLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 400
 Db 609 kttppvlidsdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispg 663

RESULT 12
 Y80123
 ID Y80123 standard; Protein; 664 AA.
 XX
 AC Y80123;
 XX

Qy 201 PKDTLMISRTPEVTCVVVDVSHEDDEVKFNWYVDGVEVHNNAKTKPREQYNSTYRVSVL 260
 Db 245 pkdtlmisrtpevtcvvdvshedpevkfnwyvdgvevhnnaatkpreeqynstyyvsvl 304
 Qy 261 TVLHQDWLNGKEYKCKVSNKALPAPIETISKAKGQPREPOVYITLPPSRDELTKNQVSLT 320
 Db 305 tvlhdqdwlngkeykckvsnkalpapietiskakgqprepqvtytlppsrdeltnqvsit 364
 Qy 321 CLVKGFPYSDIAVESNGQPENNYKTTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCS 380
 Db 365 clvkgfypsdiavewesngqpennkttppvldsdgsfflyslkltvdksrwqognvfscs 424
 Qy 381 VMHEALHNYTKQKSLSPG 400
 Db 425 vmhealhnytkqkslspsg 444

RESULT 14

W1604 ID W71604 standard; Protein; 664 AA.

XX AC W71604;

XX DT 19-NOV-1998 (first entry)

XX DE Rat neuturin receptor alpha/Fc sequence (Iff2a) fusion protein.

XX KW Human; neuturin receptor alpha; NTNR-alpha; variant; chimeric;
 fusion protein; immunoadhesion; ret-expressing cell; neurological;
 renal; haematological disease.

XX OS Synthetic.

XX OS Homo sapiens.

XX OS Rattus sp.

XX PN WO9836072-A1.

XX PD 20-AUG-1998.

XX PF 17-FEB-1998; 98WO-US031179.

XX PR 24-OCT-1997; 97US-0957063.

XX PR 18-FEB-1997; 97US-0802805.

XX PR 09-JUN-1997; 97US-0871913.

XX PA (GETH) GENENTECH INC.

XX PI Hynes MA, Klein RD, Rosenthal A;

XX PI WPI; 1998-467175/40.

XX DR N-PSDB; V58007.

XX PT New polypeptide(s) based on human neuturin receptor alpha and

XX PT related nucleic acid - useful for increasing survival of

XX PT ret-expressing cells for treating e.g. neurological, renal and

XX PT haematological diseases

XX PS Example 3; Page 84-86; 116pp; English.

XX CC The present sequence represents rat neuturin receptor alpha
 CC (NTNR-alpha)/Fc sequence (Iff2a) fusion protein, from an example of the
 CC present invention. NTNR-alpha proteins can be used: (a) to identify
 CC molecules that bind specifically to it (potential agonists and
 CC antagonists) and to purify such compounds; (b) to modulate response of
 CC cells to neuturin (NTN); (c) to increase survival of Ret-expressing
 CC cells or to activate Ret on the surface of cells (soluble glial derived
 CC neurotrophic factor receptor (GDNFR alpha) may be used the same way);
 CC (d) to increase the half-life of cognate ligands (especially NTN); (e)
 CC diagnostically to determine serum levels of its ligands; and (f) as
 CC animal feed additive or molecular weight marker. NTNR-alpha, its genes,
 CC (ant)agonists and antisense nucleic acids, are useful in vivo or ex vivo
 CC for treating conditions related to abnormal NTN activity or response,
 CC particularly neurological (central or enteric), renal or haematopoietic

CC (spleen) diseases or injuries. Ab may be agonists or antagonists for
 CC therapeutic use (e.g. as antagonists to treat excessive/unwanted NTNR-
 CC alpha expression, e.g. in some tumours), also reagents for immunosassay
 CC and affinity purification.

XX SQ Sequence 664 AA;

Query Match 55.7%; Score 1248.5; DB 19; Length 664;

Best Local Similarity 62.7%; Pred. No. 3.4e-71;

Matches 261; Conservative 21; Mismatches 61; Indels 73; Gaps 11;

Qy 32 HCTAKWTKVACPCP-DHY-----YTDSWHTSDECLYSPVCKELOVY 72
 Db 274 ncrasyrtits-cpadnyqacclgsyagmigfdmtpnyvdsnptg--ivvspwc----- 323

Qy 73 KQECNRTNHRVCEKE-----GRYLEIEFCLKHRSCTPGFGVWQAGT 114

Db 324 --ncrgsgnmeeceekflrdfenpcrlrnaiaqafngtdvmspkpslp----atqapr 377

Qy 115 PERNTVCKRCPDGFFSNETSSKAPCRKHTNCSVFCLLLLTQKGNATHD-----NIC 164

Db 378 vekt---pslpddlsdstslgtswittctsiqegll-----kannskelsmofltttnis 430

Qy 165 SGNSESTQKVDKTHTCPPCPAPELIGGFSVFLFPPKPKDTLMISRTPEVTCVVVDVSHED 224

Db 431 pg---srdpvdkthtccpapelilggpsvflfppkpkdtlmisrtpevtcvvvdvshed 487

Qy 225 PEVFNWTVDGVVHNNAKTKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPA 284

Db 488 pevfnwvyvdgvevhnnaatkpreeqynstyrsvsvltvlhqdwlngkeykckvsnkalpa 547

Qy 285 PIEKTISKAKGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFPYSDIAVESNGQPENN 344

Db 548 piektiiskakgqprepqvtytlppsrdektknqvslctclvkgfypsdiavewesngqpenn 607

Qy 345 YKTTPPVLDSGDSFFLYSLKLTVDKSRWQOGNVFSCVMHEALHNYTKQKSLSPG 400

Db 608 ykttppvlldsgdsfflyskltvdksrwqognvfscsvmhealhnytkqkslspsg 663

RESULT 15

Y80124

ID Y80124 standard; Protein; 664 AA.

XX AC Y80124;

XX DT 19-MAY-2000 (first entry)

XX DE Rat NTNR alpha and IgG fusion protein SEQ ID NO:18.

XX KW Human; neuturin receptor alpha; NTNR alpha; splenic haematopoiesis;

XX KW anaemia; thrombocytopaenia; hypoplasia; haemorrhage.

XX OS Rattus sp.

XX OS Synthetic.

XX PN US6025157-A.

XX PD 15-FEB-2000.

XX PF 24-OCT-1997; 97US-0957063.

XX PR 18-FEB-1997; 97US-0038839.

XX PR 09-JUN-1997; 97US-0049818.

XX PI (GETH) GENENTECH INC.

XX PI Hynes MA, Rosenthal A, Klein RD;

XX PI WPI; 2000-181808/16.

XX DR N-PSDB; 291460.

XX XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:20:07 ; Search time 135.68 seconds
(without alignments)
350.725 Million cell updates/sec

Title: US-09-389-782A-6
Perfect score: 2274
Sequence: 1 ETPPKYLHYDEETSHQLL.....VMHEALHNHYTKLSLSLSPG 406

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_invertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1043	45.9	372	4 Q9UHP4	Q9uhp4 homo sapien
2	1043	45.9	401	4 O00300	O00300 mus sapien
3	922.5	40.6	401	11 O08712	O08712 mus musculus
4	909	40.0	401	11 O08727	O08727 rattus norv
5	819.5	36.0	437	11 Q9R1A4	Q9ria4 mus musculus
6	433.5	19.1	300	4 Q95407	Q95407 homo sapien
7	405.5	17.8	302	13 Q9PUS0	Q9pus0 salvelinus
8	343.5	15.1	439	4 Q16042	Q16042 homo sapien
9	342	15.0	459	11 Q82327	Q82327 mus musculus
10	335	14.7	482	11 O88734	O88734 mus musculus
11	294	12.9	655	4 Q75509	Q75509 homo sapien
12	277.5	12.2	384	4 Q9UP60	Q9up60 homo sapien
13	275.5	12.1	684	13 Q90544	Q90544 ginglymosto
14	273.5	12.0	416	4 Q9NPP6	Q9npp6 homo sapien
15	253.5	11.1	625	11 Q35305	Q35305 mus musculus
16	252.5	11.1	616	4 Q9Y6Q6	Q9y6q6 homo sapien
17	241	10.6	349	12 Q57099	Q57099 monkeypox v
18	239	10.5	349	12 Q57291	Q57291 monkeypox v
19	239	10.5	349	12 Q57100	Q57100 monkeypox v

20	239	10.5	349	12	Q57101	Q57101 monkeypox v
21	239	10.5	349	12	Q57102	Q57102 monkeypox v
22	236	10.4	348	12	Q57112	Q57112 variola vir
23	236	10.4	348	12	Q85407	Q85407 variola vir
24	235.5	10.4	348	12	Q57277	Q57277 monkeypox v
25	235.5	10.4	348	12	Q57103	Q57103 monkeypox v
26	235.5	10.4	348	12	Q57108	Q57108 monkeypox v
27	233	10.2	349	12	Q57110	Q57110 variola vir
28	233	10.2	349	12	Q57111	Q57111 variola vir
29	233	10.2	349	12	Q89118	Q89118 variola vir
30	233	10.2	349	12	Q89098	Q89098 variola vir
31	228	10.0	349	12	Q57109	Q57109 variola vir
32	227.5	10.0	349	12	Q57284	Q57284 camelopox vi
33	227.5	10.0	349	12	Q57098	Q57098 camelopox vi
34	224	9.9	350	12	Q57116	Q57116 cowpox viru
35	223.5	9.8	349	12	Q57097	Q57097 camelopox vi
36	223.5	9.8	355	12	Q85308	Q85308 cowpox viru
37	221.5	9.7	349	12	Q57305	Q57305 cowpox viru
38	219.5	9.7	350	12	Q57123	Q57123 cowpox viru
39	218.5	9.6	326	12	Q57120	Q57120 cowpox viru
40	218.5	9.6	326	12	Q57122	Q57122 cowpox viru
41	218.5	9.6	351	12	Q73559	Q73559 cowpox viru
42	218.5	9.6	360	12	Q57118	Q57118 cowpox viru
43	217.5	9.6	351	12	Q57117	Q57117 cowpox viru
44	213.5	9.4	347	12	Q57115	Q57115 cowpox viru
45	211.5	9.3	316	12	Q57092	Q57092 ectromelia

ALIGNMENTS

RESULT 1
Q9UHP4 PRELIMINARY; PRT; 372 AA.
ID Q9UHP4
AC Q9UHP4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE OSTEOPTOTERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA He 2.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 31:680-684(1999).
DR EMBL; AF134187; AAF20168.1;
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368;
DR PFAM; PF00020; TNFR_c6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
FT NON_TER 1
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 QRC64;

Query Match 45.9%; Score 1043; DB 4; Length 372;
Best Local Similarity 99.4%; Pred. No. 1.5e-83;
Matches 179; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPPKYLHYDEETSHQLLCKDCPPGTYLKQHCTAKWTKVCACPDHYTDSWHTSDECL 60
Db 1 ETPPKYLHYDEETSHQLLCKDCPPGTYLKQHCTAKWTKVCACPDHYTDSWHTSDECL 60
QY 61 YCSPVKELQYVQKQECNRNHNRYVCECKEGRYLEIEFCLKHRSCPPGFGVYVQAGTPERNTV 120
Db 61 YCSPVKELQYVQKQECNRNHNRYVCECKEGRYLEIEFCLKHRSCPPGFGVYVQAGTPERNTV 120
QY 121 CKRCPPGFTSNSTSSKAPCRKHTNCSVFGLLLTQKNATHDNICSGNSESTQKCGIDVTV 180
|||||

Db 121 CKRCPDFFSNSTSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQKCGIDVTL 180

RESULT 2

000300 PRELIMINARY; PRT; 401 AA.

AC 000300; 060236;

DT 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMBlrel. 04, Last annotation update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)

DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11B).

GN TNFRSF11B OR OPG OR OCIF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RX MEDLINE=97262071; PubMed=9108485;

RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,

RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,

RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,

RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,

RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,

RA Suggs S., Boyle W.J.;

RA "Osteoprotegerin: a novel secreted protein involved in the regulation

RT of bone density.";

RL Cell 89:309-319(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=LUNG FIBROBLAST;

RX MEDLINE=98151033; PubMed=9492069;

RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,

RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,

RA Tsuda E., Morinaga T., Higashio K.;

RA "Identity of osteoclastogenesis inhibitory factor (OCIF) and

RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits

RT osteoclastogenesis in vitro.";

RL Endocrinology 139:1329-1337(1998).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RX MEDLINE=98351569; PubMed=9688283;

RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;

RA "Cloning and characterization of the gene encoding human

RT osteoprotegerin/osteoclastogenesis-inhibitory factor.";

RL Eur. J. Biochem. 254:685-691(1998).

CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES

CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY

CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO

CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN

CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.

CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,

CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN

CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN

CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.

CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

DR EMBL; AB002146; BAA25910.1; -

DR EMBL; AB008822; BAA32076.1; -

DR EMBL; AB008821; BAA32076.1; JOINED.

DR EMBL; U94332; AAB53709.1; -

DR HSP; P25942; ICDF.

DR MIM; 602643; -

DR INTERPRO; IPR001368; -

DR PFAM; PF00020; TNFR_C6; 3.

DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

DR PROSITE; PS50050; TNFR_NGFR_2; 2.

DR PRODOM; PD000771; -; 1.

Glycoprotein; Repeat; Cytokine; Signal.

KW SIGNAL 1 21

FT CHAIN 22 401

FT DOMAIN 23 183

FT REPEAT 23 63

FT REPEAT 64 106

FT REPEAT 107 143

FT REPEAT 144 201

FT DOMAIN 306 365

FT DISULFID 41 54

FT DISULFID 44 62

FT DISULFID 65 80

FT DISULFID 83 97

FT DISULFID 87 105

FT DISULFID 118 142

FT DISULFID 145 160

FT CARBOHYD 98 98

FT CARBOHYD 152 152

FT CARBOHYD 165 165

FT CARBOHYD 178 178

FT CARBOHYD 289 289

FT CONFLICT 263 263

SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 45.9%; Score 1043; DB 4; Length 401;

Best Local Similarity 99.4%; Pred. No. 1.6e-83;

Matches 179; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPPKYLHYDEETSHQLCDKCPGTYLKQHCCTAKWTKVCAPCPDHYTDSNHTSDECL 60

Db 22 ETPPKYLHYDEETSHQLCDKCPGTYLKQHCCTAKWTKVCAPCPDHYTDSNHTSDECL 81

QY 61 YCSPVKELQYVQECNTHNRVCECKEGRYLETFECLKHKRSCPPGVQAGTPERTV 120

Db 82 YCSPVKELQYVQECNTHNRVCECKEGRYLETFECLKHKRSCPPGVQAGTPERTV 141

QY 121 CKRCPDFFSNSTSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQKCGIDVTL 180

Db 142 CKRCPDFFSNSTSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQKCGIDVTL 201

RESULT 3

008712 PRELIMINARY; PRT; 401 AA.

ID 008712; 070202;

AC 008712; 070202;

DT 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)

DE (OCIF).

GN TNFRSF11B OR OPG.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALE/C; TISSUE=KIDNEY;

RX MEDLINE=97262071; PubMed=9108485;

RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,

RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,

RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,

RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,

RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,

RA Suggs S., Boyle W.J.;

RA "Osteoprotegerin: a novel secreted protein involved in the regulation

RT of bone density.";

RL Cell 89:309-319(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129/OLA, AND NIH SWISS;

RX MEDLINE=98382527; PubMed=9714833;

RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,
RA Higashio K.;
RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
RT gene and its expression in embryogenesis.";
RL Gene 215:339-343(1998).
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,
CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND
CC PLACENTA. NOT DETECTED IN SPLEEN.
CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT
CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY
CC 15 TO DAY 17.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94331; AAB53708.1; -.
DR EMBL; AB013898; BAA28269.1; -.
DR EMBL; AB013903; BAA33388.1; -.
DR EMBL; AB013899; BAA33388.1; JOINED.
DR EMBL; AB013900; BAA33388.1; JOINED.
DR EMBL; AB013901; BAA33388.1; JOINED.
DR EMBL; AB013902; BAA33388.1; JOINED.
DR HSSP; P25942; ICDF.
DR MGD; MGI:109587; Opg.
DR INTERPRO; IPR000488; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 201 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT VARIANT 138 138 SWISS.
FT VARIANT 161 161 R -> P (IN STRAINS 129/OLA AND NIH
FT VARIANT 165 165 I -> R (IN STRAINS 129/OLA AND NIH
FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH
FT VARIANT 288 288 S -> A (IN STRAINS 129/OLA AND NIH
FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH
FT SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;

Query Match 40.6%; Score 922.5; DB 11; Length 401;
Best Local Similarity 48.3%; Pred. No. 6e-73;
Matches 196; Conservative 35; Mismatches 102; Indels 73; Gaps 9;
QY 1 ETTPPKYLHYDEETSHQLLCKDCPPTGLKQHCTAKWKTCAPCPDHYYTDSWHTSDECL 60

Db 22 ETLPKYLHYDETPGTHQLLCKDCAPCTYKQHCTVRRKTLVPCPDHSTDSWHTSDEC 81
QY 61 YCSPVKELQYKQECNRTNHRVCECKEGRYLEIEFCLKHRSCPPGFGVVGAGTPERTIV 120
Db 82 YCSPVKELQYKQECNRTNHRVCECKEGRYLEIEFCLKHRSCPPGFGVVGAGTPERTIV 141
QY 121 CKRCPDGFSTSSKAPCKRKHNTCSVFGLLLTQKNATHDNCISGNSSESTQKCGIDVT 180
Db 142 CKKCPDGFSTSSKAPCKRKHNTCSVFGLLLTQKNATHDNCISGNSSESTQKCGIDVT- 200
QY 181 DKTHTCPPCAPPELLGSPSVFLPPPKPDTLMISRTPVTVVVDVSHEDPEYKFNMYVD 240
Db 201 -----LCEEAFFRAVPTK-----IIPNWSLVLD---SLPGTKVN--AE 235
QY 241 GVEVHNNAKTPREEQYNSTYRVSVLTVLHQDLNKGKEY-----CKKVS----- 284
Db 236 SVE-----RIKRHSSEOETFOLLKWKHQRDOEMVKIIQIDIDLCSESVQRHLGHSNLT 291
QY 285 -----NKALPAPIEKTISKAKGPPEOVYTL-----PPSRDELTKNQVSLTC 327
Db 292 TEQLLALMESLPQKKISPEIERTRKTSSEQLLKLKLSLWRKNGDQDTLKGMLTAKH 351
QY 328 LVKGFYPSDIAVWESNGOPENNYKTTTPPVLDSDGSFFLYSKLTVD 373
Db 352 LKTSHEPKVT-----HSLRKTMRFLHSFTWRYLQKLFLE 387
RESULT 4
O08727 PRELIMINARY; PRT; 401 AA.
ID O08727
AC O08727
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF).
GN TNFRSF11B OR OPG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RT *Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density.*;
RL Cell 89:309-319(1997).
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94330; AAB53707.1; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21 BY SIMILARITY.

```
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 201 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
SQ SEQUENCE 401 AA; 46192 MW; FECA31F1D4E573A CRC64;

Query Match 40.0%; Score 909; DB 11; Length 401;
Best Local Similarity 85.6%; Pred. No. 9.1e-72;
Matches 154; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 ETPFPKYLHYDEETSHOLLCDKCPGTYLKHCHTAKWKTVCAPCPDHVYDTSWHTSDECL 60
DB 22 ETPFPKYLHYDETPGROLLCDKCAPGYLYLKHCHTVRKTLVCPDYSDYDTSWHTSDECV 81

QY 61 YCSPVKELQVQKQECNTHNRVCEKRGYLETEFCCLKHRSCTPPGGVQVQAGTPERNV 120
DB 82 YCSPVKELQVQKQECNTHNRVCEKRGYLETEFCCLKHRSCTPPGGVQVQAGTPERNV 141

QY 121 CKRCPDFFSNETSKAPCRKHTNCSVFGILLITOKGNATHNICSNGSESTOKCGIDVTY 180
DB 142 CKRCPDFFSNETSKAPCRKHTNCSVFGILLITOKGNATHNICSNGSESTOKCGIDVTY 201

RESULT 5
Q9RI4 PRELIMINARY; PRT; 437 AA.
AC Q9RI4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE GAMMA1 HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFV).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD04243.1; -
DR HSP; P01842; 7FAB.
DR INTERPRO; IPR003006; -.
DR FRAM; PF00047; 1g; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR NON_TER 1 1
FT NON_TER 437 437
FT SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 36.0%; Score 819.5; DB 11; Length 437;
Best Local Similarity 62.8%; Pred. No. 7.1e-64;
Matches 140; Conservative 43; Mismatches 35; Indels 5; Gaps 2;

QY 186 CPP--CRAPELLGSGVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNYYDGVGE 243
DB 186 CPP--CRAPELLGSGVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNYYDGVGE 243
```

```
DB 217 CKPICICTVPEV---SSVFIFFPKPKDVLITLTPKVTVCVVVDISKDDPEVQFSWFVDVDE 273
QY 244 VHNAKTPREOYNSTYRVYVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAKGQP 303
DB 274 VHTAQTPREOYNSTYRVYVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAKGQP 333
QY 304 REPOVYTLPPSRDELTKNOVSLTLVKRGFYPSDIAVWESNGOPENNYKTPPVLDSDGS 363
DB 334 KAPQVYTIPTPPKEQMAKDKVSLTCMTIDFFPEDITVEMQWNGQPAENYKNTQPIMDTDS 393
QY 364 FFYLSKLTVDKSRWQOQNVFSCSVHHEALHNHTYKQSLSPG 406
DB 394 YFYSKLVNOKSNWEAGNTFTCSVLHGLHNHHTKLNLSHPG 436

RESULT 6
Q95407 PRELIMINARY; PRT; 300 AA.
AC Q95407;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA Goddard A.D., Botsstein D., Ashkenazi A.;
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
RT colon cancer.";
RL Nature 396:699-703(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=BLOOD;
RX MEDLINE=99253915; PubMed=10318773;
RT Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT "A newly identified member of tumor necrosis factor receptor
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL J. Biol. Chem. 274:13733-13736(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=PANCREAS;
RX MEDLINE=20122600; PubMed=10655513;
RT Bai C., Connolly B., Metzger M.L., Hilliard C.A., Liu X., Sandig V.,
RT Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RT cluster.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR EMBL; AF104419; AAD03056.1; -
DR EMBL; AF134240; AAD29688.1; -
DR EMBL; AF217796; AAF35244.1; -
DR EMBL; AF217793; AAF33685.1; -
DR EMBL; AF217794; AAF33686.1; -
DR HSP; P25942; 1CDF.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR FRAM; PF00020; TNFR_c6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;
```

Query Match 19.1%; Score 433.5; DB 4; Length 300;
 Best Local Similarity 39.1%; Pred. No. 3.1e-30;
 Matches 72; Conservative 32; Mismatches 69; Indels 11; Gaps 1;

QY 5 PKYHYDEETSHQLLCKDPCPGTYLKQHTAKWTVCAPCPDHYHYDSDWTSDECLYCSP 64
 Db 34 PTPYWRDAETGERLVAQCPGPGTYVQPCRPDRPTTCGPPRHYTFQWNYLERCRYCNV 93
 QY 65 VKELQYVQKQENRTHNRYCECKEGRYLEFCLKHRSCPPGFGVVOAGTPERTNVCCKRC 124
 Db 94 LCGEREERARACHATHNRACRTGTFHAHAGFCLHASCPPGAGVATGPTPSONTOCQPC 153
 QY 125 PDGFFNETSSKAPCRKHTNCVFGLLLTQKGNATHDNCISG-----NSESTQK 173
 Db 154 PPGTFASSSSSQCPQHRNCTALGLALNVPGSSSHDTLCTSGTGPLSTRVPGAECER 213
 QY 174 CGTD 177
 Db 214 AVID 217

RESULT 7
 Q9PUS0 PRELIMINARY; PRT; 302 AA.

AC Q9PUS0;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE DECOY RECEPTOR.
 OS Salvelinus fontinalis (Brook trout).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
 OX NCBI_TaxID=8038;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bobe J., Goetz F.W.;
 RT "A tumor necrosis factor receptor homolog is up-regulated in the brook trout (Salvelinus fontinalis) ovary at the completion of ovulation";
 RL Biol. Reprod. 0:0-0(1999).
 DR EMBL; AF156738; AD56428.1; -.
 DR HSSP; P19438; IEXT.
 DR INTERPRO; IPR000561; -.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 1.
 KW Receptor.
 SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 17.8%; Score 405.5; DB 13; Length 302;
 Best Local Similarity 45.2%; Pred. No. 8.9e-28;
 Matches 71; Conservative 30; Mismatches 53; Indels 3; Gaps 2;

QY 11 DEETSHQLLCKDPCPGTYLKQHTAKWTVCAPCPDHYHYDSDWTSDECLYCSPCKELQ 70
 Db 27 DRYSGLSIVCDRCPPGTYLRAPCSARKSDCAECPNGAYTEFWNHISKLCURCS-MCAENQ 85

QY 71 YVQKQENRTHNRYCECKEGRYLEFCLKHRSCPPGFGVVOAGTPERTNVCCKRCPDGF 128
 Db 86 VVKQESPSNNCECKEGYFNKKEACIKHKECPGANTGTGPHQDTECVQOAGF 145

QY 129 FSNETSAPCRKHTNCVFGLLLTQKGNATHDNCIS 165
 Db 146 YSEVSSAKATCLAQSNCKVGLRWLVKGDWHNTLCA 182

RESULT 8
 Q16042 PRELIMINARY; PRT; 439 AA.
 ID Q16042
 AC Q16042;

DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91370690; PubMed=1966549;
 RA Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R., Brockhaus M., Lesslauer W.;
 RT "Two human TNF receptors have similar extracellular, but distinct intracellular, domain sequences";
 RL Cytokine 2:231-237(1990).
 DR EMBL; S63368; AAB19824.1; -.
 DR HSSP; P25942; ICDP.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 3.
 DR PRODOM; PD000771; -.
 SQ SEQUENCE 439 AA; 46090 MW; FEBCE329CC67FF6 CRC64;

Query Match 15.1%; Score 343.5; DB 4; Length 439;
 Best Local Similarity 27.3%; Pred. No. 3.8e-22;
 Matches 110; Conservative 44; Mismatches 162; Indels 87; Gaps 16;

QY 9 HYDEETSHQLLCKDPCPGTYLKQHTAKWTVCAPCPDHYHYDSDWTSDECLYCSPVCK 68
 Db 23 YIDQ--TAQMCCKSCSPGQAHKVFCTKSTVTCDSCESTYTLNWNVPCLSCGRCSS 80

QY 69 LQYVQKQENRTHNRYCECKEGRYLEI-----EFLKHRSCPPGFGVVOAGTPERTNVC 122
 Db 81 DQVETQACTREQNRICTCRPHQICNVVAI-----PGNASMDAVCTSTSPTRSMAPGAHLQ 140

QY 123 RCPDGFESNETSSKAPCRKHTNCVFGLLLTQKGNATHDNCISNSEQTKC-----GI 176
 Db 141 PCAPGTFSTSTSDICRPHQICNVVAI-----PGNASMDAVCTSTSPTRSMAPGAHLQ 196

QY 177 DVTVDKTHCTCP---PCPAPE---LLGPGSVFLFPKPKDT-----LMSRTP--- 217
 Db 197 PVSTRSQHTQPTPEPSTAPSTSLPMGPS-----PPAEGSTGDFALPVLIVGTALGLL 252

QY 218 ---EVTGVVVDVSHEDP-----EVKFNWYVDGVEVHNK-TKPREQYNTYRVSVLT 268
 Db 253 IIGVWNCVIMTVKKKPLCLQREAK-----VPHLPADKARGTQGPQOQ-----LIT 300

QY 269 LHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTK----- 320
 Db 301 APSSSSSLESSASALDRRAP-----TRNQAPQAPGEASGAGARASTGSSDSPG 351

QY 321 ---NOVSLTCLVKGFPSPDIAVEWESN-----GOPENNYKTP 355
 Db 352 GHGTQVNTCIVNVCSSSDHSSQSSQASSTMGDTDSSPESP 394

RESULT 9
 Q62327 PRELIMINARY; PRT; 459 AA.
 ID Q62327
 AC Q62327;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

```
RC STRAIN=NOD;
RA MEDLINE=95178848; PubMed=7873884;
RX Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
RL gene.";
DR Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.
DR HSP; P19438; INCF.
DR INTERPRO: IPR001368; -.
DR PFAM: PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
FT NON_TER 1
FT VARIANT 87 87 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match 15.08; Score 342; DB 11; Length 459;
Best Local Similarity 25.84; Pred. No. 5.4e-22;
Matches 103; Conservative 52; Mismatches 166; Indels 78; Gaps 13;

Qy 9 HYDETSHQLLCKPCPGTYLKQHCCTAKWKTVCAPCPDHYVYDTSWHTSDECLY 68
Db 31 YDRKA--QMCCAKCPGQYVYKHCNKTSDIVCADCEASMTQVWNOFRTCLS 88
Qy 69 LOYVQKQECNRTHNRVCEKGRYLEIEF-----CLKHRSPPGFGVVGAGT 121
Db 89 DQVETRACTQONRVCAACEAGRYCALKTHSGRCQCMRLSKCGPGFVASSRA 148
Qy 122 KRCPDGPFSTSSKAPCRKHTNCSVFGLLLTQKGNATHNICNSSESTOKCGIDVT 180
Db 149 KACAPGTFSDTSTDVCRPHRISILAI-----PGNASTDAVCA--PESPTLSAIPRTLY 202
Qy 181 ----DKHTCTCP-----PCPAPELL-----GGPSV-----FLFP 204
Db 203 VSQPEPTRSQPLDQEPGSPQPSILTSGLSTPIEQSTKGISLPIGLIVGTVSLGLML 262
Qy 205 PKPKDTLMISTPEVTCVVDVSHEDPEVKFNWYVDGVGVHNAKTKPREQYNSTYRVVS 264
Db 263 GLVNCFLVQRKKKPSCLQRDA--KVHPVDEKSDAVGL-----EQQH----- 304
Qy 265 VLTVLHQLDNLGKEYCKVSKNALPA---PIEKTISKAKGQPREPQVYTLPPSRDELTKN 321
Db 305 LLTAPSSSSSSLSASSAGDRAPPQGHGPOARVMAEAGQSGEASRRISDSHSGHGT 364
Qy 322 QVSLTCLVKGFPYSDIAVEWESN-----GOPENNYKTP 355
Db 365 HNVNCTIVNVCSSDHSQCSQASATVGDPAKPSASP 403

RESULT 10
O88734 PRELIMINARY; PRT; 482 AA.
ID O88734
AC O88734
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE P80 TNF-ALPHA RECEPTOR.
GN TNFR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hurler B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;
RT "The Mouse Tumor Necrosis Factor Receptor Gene: Genomic Structure and
RL Characterization of the two Transcripts.";
RL Genomics 0:0-0(0).
```

```
DR EMBL; Y14619; CAA74969.1; -.
DR EMBL; Y14620; CAA74969.1; JOINED.
DR EMBL; Y14621; CAA74969.1; JOINED.
DR EMBL; Y14622; CAA74969.1; JOINED.
DR EMBL; Y14623; CAA74969.1; JOINED.
DR EMBL; Y14629; CAA74969.1; JOINED.
DR HSP; P19438; INCF.
DR INTERPRO: IPR001368; -.
DR PFAM: PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
DR PRODOM; PD000771; -.
SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;

Query Match 14.78; Score 335; DB 11; Length 482;
Best Local Similarity 25.78; Pred. No. 2.3e-21;
Matches 105; Conservative 49; Mismatches 166; Indels 88; Gaps 14;

Qy 9 HYDETSHQLLCKPCPGTYLKQHCCTAKWKTVCAPCPDHYVYDTSWHTSDECLY 61
Db 46 YDRKA--QMCCAKCPGQYVYKHCNKTSDIVCADCEASMTQVWNOFRTCLS 103
Qy 62 CSPVCKELQYVQKQECNRTHNRVCEKGRYLEIEF-----CLKHRSPPGFGVVGAGT 114
Db 104 CSSSCSTDQVETRACTQONRVCAACEAGRYCALKTHSGRCQCMRLSKCGPGFVASSRA 163
Qy 115 PERNTVCKRCPDGFSTSSKAPCRKHTNCSVFGLLLTQKGNATHNICNSSESTOKC 174
Db 164 PNGVLKACAPGTFSDTSTDVCRPHRISILAI-----PGNASTDAVCA--PESPTLS 217
Qy 175 GIDVTV-----DKHTCTCP-----PCPAPELL-----GGPSV----- 200
Db 218 AIPRTLVQPEPTRSQPLDQEPGSPQPSILTSGLSTPIEQSTKGISLPIGLIVGTV 277
Qy 201 ---FLFPKPKDTLMISTPEVTCVVDV--VSHEDPEVKFNWYVDGVGVHNAKTKPREEQ 255
Db 278 SLGLMLGLVNCFLVQRKKKPSCLQRDAKVPHV-PDEKSDAVGLEQHLHTTAPSSSS 336
Qy 256 YNSYRVVSVLTVLHQLDNLGKEYCKVSKNALPA---PIEKTISKAKGQPREPQVYTL 312
Db 337 SSSL-----ESSASAGDRAPPQGHGPOARVMAEAGQSGEASRRIS 378
Qy 313 PSRDELTKNOVSLTKLVKGFPYSDIAVEWESN-----GOPENNYKTP 355
Db 379 DSSHGSHGTHVNTCTIVNVCSSDHSQCSQASATVGDPAKPSASP 426

RESULT 11
O75509 PRELIMINARY; PRT; 655 AA.
ID O75509
AC O75509
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TNFR-RELATED DEATH RECEPTOR-6 (DJ181J13.1) (DR6 (TNFR-RELATED DEATH
DE RECEPTOR-6)).
GN DR6 OR DJ181J13.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pan G., Bauer J.H., Haridas V., Wang S., Liu D., Ni J., Yu G.,
RA Vincenz C., Aggarwal B.B., Dixit V.M.;
RT "Identification and functional characterization of DR6, a novel death
RT domain-containing TNF receptor.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Parker A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AF068868; AAC34583.1; -.
DR EMBL; AL096801; CAB75692.1; -.
DR HSSP; P07174; INGR.
DR INTERPRO; IPR000488; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 4.
DR PFAM; PF00531; death; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00117; DEATH_DOMAIN; 1.
DR PROSITE; PS00050; TNFR_NGFR_2; 1.
DR PRODOM; PD000771; -.
KW Receptor.
SQ SEQUENCE 655 AA; 71844 MW; 48939391C4852A33 CRC64;

Query Match 12.9%; Score 294; DB 4; Length 655;
Best Local Similarity 27.8%; Pred. No. 1.3e-17;
Matches 87; Conservative 44; Mismatches 134; Indels 48; Gaps 11;

QY 7 YLHYDETHSHLLCDKCPGTYLKQHTAKWTKVACAPCDHYHYTDSWHTSDECLYCSPVC 66
DB 54 YRHVDRTAGVLTCDKCPAGTYSEHCTNTSLRVCSGPCVGTFRHENGIEKCHDCSQPC 113
QY 67 KELQYVQECNRTHNRVCECKEGRYLEIEFCLKHSRCPGPGVGVQAGTPERNVCKRCPD 126
DB 114 PPMIEKLPALCAALTDRECTCPGFMFSQSNATCAPHTVCPVGWGVKRGKGTETEDVRCKQCAR 173
QY 127 GFPSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNICS---CNSBSTQKCGIDVTVDYKT 193
DB 174 GTESDVPSSVMKCAVTKDCLSLQNLVVIKPGTKETDNVCGTLPFSFSSSTS----- 222
QY 184 HTCPCPAPPELLGGSPVFLPPPKPKDTLMISRTPEVTCVVVDVSHED-----PEVKF 235
DB 223 -----PSP-----GTAIF---PRP-EMETHPEVPSSTYVPKGMNSESNSASVRKV-L 267
QY 236 NWTVDGVEVHNATKPREQYNTYRVSVVSVLTVLVHODWLNKGYCKVCSNKAALPAPI--- 292
DB 268 SSTQEGTVPDNTSSARGKEDVNKT--LPNLQVNVNHQ---QGPHRHIL--KLLPSMEATG 320
QY 293 -EKTISKAKGQPR 304
DB 321 GEKSTPIKPKR 333

RESULT 12
Q9UP60 PRELIMINARY; PRT; 384 AA.
AC Q9UP60;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SNC73 PROTEIN.
GN SNC73.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;
RT "Identification and characterization of SNC73, a gene which is down-
RT regulated in colorectal cancer.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067420; AAC19365.1; -.
DR HSSP; P01825; 7FAB.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;

Query Match 12.2%; Score 277.5; DB 4; Length 384;
Best Local Similarity 28.2%; Pred. No. 2e-16;
```

```
Matches 96; Conservative 35; Mismatches 138; Indels 71; Gaps 14;

QY 115 PERNTCKRCPCDGFESNE-----TSSKAPCRKHTNCSVFGLLLTOK 155
DB 49 PDGNVVIACLVQGFQFQEPISVTSWESGGQGVTAARNFPSPQDASGLYTTSSQTLTPATQ- 107
QY 156 GNATHDNICSGNSESTOKCGI-----DVTVDKTHCTCPAPPELLGGSPVFLFPKP 207
DB 108 -----CLAGKSVT--CHVKHTNPQDVT-----PCVPSPPTPSPST-PPTP 149
QY 208 K-----DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPR 252
DB 150 SPSCCHPRLSLHRPALEDLLLGSEANLTCLTGL-RDASGVTTETWTPSSGK--SAVQGP 206
QY 253 EEOYNSTYRVSVLTVLVHODWLNKGYCKVCSNKAALPAPIEKTISKAKGQPREPOVYTL 312
DB 207 ERLDGCYSVSSVLPGCAEPWNHGKTFCTAAYPESTKPLTATLSKS-GNTFRPEVHLLP 265
QY 313 PSRDELTKNQ-VSLTCLVKGFPYSDIAVEWESNQ--PENNYKTTTPVLD--SDG--SFFL 366
DB 266 PPEELALNELVLTCLARGFSPKDLVLRWLQSOELPREKYLTWASRQSPSQGTTTFAV 325
QY 367 YSKLTVDKSRMQGNVFCSCVMHEALHNHYTQKSLSLSPG 406
DB 326 TSLRVAADWKKGDTFCMVHGHEALPLAFTQKTIDRLAG 365

RESULT 13
Q90544 PRELIMINARY; PRT; 684 AA.
AC Q90544;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]
RP SEQUENCE FROM N.A.
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks.";
RL Nature 374:168-173(1995).
DR EMBL; U18701; AAB48195.1; -.
DR HSSP; P01857; IFC1.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; ig; 6.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Signal.
FT SIGNAL.
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 12.1%; Score 275.5; DB 13; Length 684;
Best Local Similarity 26.3%; Pred. No. 5.8e-16;
Matches 101; Conservative 53; Mismatches 161; Indels 69; Gaps 17;

QY 51 DSWHTSDECLYCS-----PVCKELQYVQECNRTHNRVCECKEGRYLEIEFCLKHS 102
DB 315 EEMQSGVE-YTCSAKQDQSSTPVVKTRKARVEPTKPHRL----- 354
QY 103 CPFGFGVGVQAGTPERNVCKRCPDGFPSNETS---SKAPCRKHTNCSVFGLLLTQKNAT 159
DB 355 LPPSPBIEQTSSTATLTCLIR---GFYPDKVSVSWQDDVSVSANVTNFTALEQ--DLT 409
QY 160 HDNICSNSESTQ-KCGIDVTVDKTHTCPP-----CPAPPELLGGSPVFLF 203
```

```

Db 410 FSTRLLNLTAWEKSGAKYCTASH--PPSQSTVKRVRIRNQKDCRQTDI-----SVSLL 463
QY 204 PPKPKDLMISRPEVTCVVVDVSHEDPE-VKFNWYVDGVEVHNAKTKPREEQYNSYRV 262
Db 464 KP-PFEEIWTQQTATVCEIV---YSLENIKFWQNGVERKKGVTQNPWESGSKSTI 519
QY 263 VSVLTVLHQDLNGKEYCKVKSNKALPAPIETKISKAK-GOPREPQVYTLPPSRDEL-TK 320
Db 520 VSKLVNASEWSDSTEYCLVDESELTVPVKASIRKANVSQHPKPKVYLLHPSTDEIDTE 579
QY 321 NOVSLATCLVKGFPDIAEVESNGQ-PENNYKTPPVLDSDGFFLYSKLTVDKSRWQQ 379
Db 580 NSATLMCLATNFPAEYVGMANDTLIDSGYRQTVDSKSGSFVTDRLRLTAAEWNS 639
QY 380 GNVFSCVMHEALHN---HYTKS 400
Db 640 DTTYSLVGHPSLRDLIRSTNKS 663

RESULT 14
ID Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Escivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA Auffray C., Ansoere W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RT "The European IMAGE consortium for integrated Molecular analysis of
human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389578; CAB97534.1; -.
FT NON_TPR 1
FT SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 12.0%; Score 273.5; DB 4; Length 416;
Best Local Similarity 27.5%; Pred. No. 4.8e-16;
Matches 96; Conservative 37; Mismatches 137; Indels 79; Gaps 15;

QY 104 PPGGVVQAGTPERTVCKRC-PDGFFSNE-----TSSKAPCRKHT 143
Db 82 KPVEPLSDSTPDGQNVVACLVQGFPPQEPPLSVTWSGQNVTFARPPSQDASGLYT 141
QY 144 NCSVEGLLTQGNATHDNICSGNSESTQKGI-----DVTVDKTHCTPCCPAPPELL 195
Db 142 TSSQLTLPATQ-----CPDGKSVT---CHVKHYNPSQDVTV-----PCPVP--- 180
QY 196 GGPVSVFLFPKPP-----KDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVE 243
Db 181 -----PPPCCHPRLSLRPALEDLLGSEANLCTLTGL-RDASGATFTTTPSSGK 231
QY 244 VHNKATPREEQYNSTRVSVLTVLHQDLNGLNGKEYCKVKSNKALPAPIETKISKAKQP 303
Db 232 --SAVQPPPERDLGCGYSVSVLPGCAQPNHGETFTCTAAHPELKTPLTANITKS-GNT 288
QY 304 REPOVYTLPPSRDELTKNQ-VSLTCLVKGYFVPSPDIAEVESNGO--PENNYKTPPVLD- 359
Db 289 FRPEVHLLPPPESEALNELVTLTCLAGFSPKDVLRWLGQSGELPREKYLTVASRQEP 348
QY 360 SDG--SFFLYSKLTVDKSRWQNGQNVFSCVMHEALHNHYTQKSLSLSPG 406

```

```

Db 349 SQGTTTFAVTSILRVAADWKKGDTFSCMGHEALPLAFTQKTDRLAG 397

RESULT 15
ID O35305 PRELIMINARY; PRT; 625 AA.
AC O35305;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE RECEPTOR ACTIVATOR OF NF-KAPPA-B PRECURSOR (TNF-RELATED ACTIVATION-
DE INDUCED CYTOKINE RECEPTOR) (RANK).
GN TNFRSF11A OR RANK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER EPITHELIUM;
RX MEDLINE=98032977; PubMed=9367155;
RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
RA Tometsko M.E., Roux E.R., Teepe M.C., DuBose R.F., Cosman D.,
RA Galibert L.;
RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
RT and dendritic-cell function.";
RL Nature 390:175-179(1997).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS RANKL.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; AF019046; AAB86810.1; -.
DR HSP; P25942; ICDF.
DR MGD; MGI:1314891; Tnfrsf11a.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 1.
DR PRODOM; PD000771; -.
DR RECEPTOR; Glycoprotein; Transmembrane; Signal; Repeat.
FT SIGNAL 1 30
FT CHAIN 31 625
FT DOMAIN 31 214
FT TRANSNEM 215 235
FT DOMAIN 236 625
FT DOMAIN 34 196
FT REPEAT 34 70
FT REPEAT 71 114
FT REPEAT 115 153
FT REPEAT 154 196
FT REPEAT 35 47
FT DISULFID 48 61
FT DISULFID 51 69
FT DISULFID 72 87
FT DISULFID 93 113
FT DISULFID 115 125
FT DISULFID 127 134
FT DISULFID 128 152
FT DISULFID 155 170
FT DISULFID 176 195
FT CARBOHYD 106 106
FT CARBOHYD 175 175
FT SEQUENCE 625 AA; 66621 MW; F8C1872E99511D8E CRC64;

Query Match 11.1%; Score 253.5; DB 11; Length 625;
Best Local Similarity 24.9%; Pred. No. 4.4e-14;
Matches 102; Conservative 41; Mismatches 147; Indels 119; Gaps 17;

QY 20 CDKCPGPGTYLKQHCTAKWTKVCAFCPDHYTDSWHTSDECLYCPVC---KELYVVKQEC 76

```

Db 48 CSCEPGKYLSSKCTPTSDVCLPCGPDYLDYLTWNEEDKCL-LHKVCDAGKALVAV-DPG 105
QY 77 NRTHNRVCECKEGRY--LEIEFCLKHRSPPGPGVVQAGTPERTYCKRCPDGFFSNETS 134
Db 106 NHTAPRCACACTAGYHWNDCCECRRNTECAPGEGAGHPLQLNKDVTCTPCLLGGFFSDVFS 165
QY 135 SKAPCRKHTNCVFGLLLTQKNATHDNTCSGNSESTQKCGIDVVDKTHTCPPCPAPEL 194
Db 166 STDKCKPWTNCTLLGKLEAHQGTESDVVCS--MTLRRPPKEAQAY 211
QY 195 LGGPSVLEFPKPKDPLMISRTPEVTCVVVDYSHEDPEVKFNWYVDGVEVHNAKTKPREE 254
Db 212 L--PSLIVL-----LLFISVVVAAIIFGV-----YVRKGGK----- 241
QY 255 QYNSTYRVVSVLTVLHODWLN-----GKEY---KCKVSNKA-----LPAP 291
Db 242 -----ALTANLWNWVNDACSSLSGNKSSGDRGAGSHSATSSQOEVCBGILLMTR 291
QY 292 IEKTISK-----AKQP-----REPQVYTL-----PSRDELT--KNO 322
Db 292 EEKWPEDGAGVGPVCAAGGPAEVRDSRTFTLVSEVETQGDLSRKIPTEDYTDORPSQ 351
QY 323 VSLTCLV-----KGFYPSDIAVEWESNGOPENNYKTTTPVLDSGSPF 365
Db 352 PSTGSLLLIQGSKSIPFPQEPLEVGENDLSQCFTGTSTVDSEGDF 400

Search completed: March 1, 2001, 09:20:09
Job time: 407 sec

Result No.	Query	Score	Query			DB	ID	Description
			Match	Length	Count			
1	1231	54.1	330	1	GCL_HUMAN	P01857	homo sapien	
2	1148	50.5	326	1	GC2_HUMAN	P01859	homo sapien	
3	1138.5	50.1	327	1	GC4_HUMAN	P01861	homo sapien	
4	1134.5	49.9	290	1	GC3_HUMAN	P01860	homo sapien	
5	917	40.3	323	1	GC_RABIT	P01870	oryctolagus	
6	895.5	39.4	329	1	GC2_CAVPO	P01862	cavia porcea	
7	848.5	37.3	329	1	GC3_MOUSE	P22436	mus musculus	
8	842.5	37.0	398	1	GC3M_MOUSE	P03987	mus musculus	
9	824	36.7	333	1	GC3_RAT	P02061	rattus norv	
10	817.5	35.9	326	1	GCL_RAT	P020759	rattus norv	
11	812.5	35.7	324	1	GCL_MOUSE	P01868	mus musculus	
12	812.5	35.7	393	1	GCLM_MOUSE	P01869	mus musculus	
13	805	35.4	329	1	GCC_RAT	P020762	rattus norv	
14	804	35.4	330	1	GCAA_MOUSE	P01863	mus musculus	
15	804	35.4	399	1	GCAM_MOUSE	P01865	mus musculus	
16	797	35.0	335	1	GCAB_MOUSE	P01864	mus musculus	
17	782	34.4	322	1	GCA_RAT	P020760	rattus norv	
18	779.5	34.3	336	1	GCB_MOUSE	P01866	mus musculus	
19	779.5	34.3	405	1	GCBM_MOUSE	P01867	mus musculus	
20	366.5	16.1	421	1	EPC_MOUSE	P06336	mus musculus	
21	361.5	15.9	391	1	MUCB_HUMAN	P04220	homo sapien	
22	359	15.8	454	1	MUC_HUMAN	P01871	homo sapien	
23	355.5	15.6	429	1	EPC_RAT	P01855	rattus norv	
24	354	15.6	455	1	MUC_MOUSE	P01872	mus musculus	
25	352	15.5	428	1	EPC_HUMAN	P01854	homo sapien	
26	349	15.3	476	1	MUCM_MOUSE	P01873	mus musculus	
27	346.5	15.2	458	1	MUC_RABIT	P03988	oryctolagus	
28	345.5	15.2	461	1	TNR2_HUMAN	P02033	homo sapien	
29	341.5	15.0	479	1	MUCM_RABIT	P04221	oryctolagus	
30	337	14.8	474	1	TNR2_MOUSE	P25119	mus musculus	
31	335.5	14.8	450	1	MUC_CANFA	P01874	canis famill	
32	332.5	14.5	454	1	MUC_MESAU	P06337	mesocricetu	
33	329.5	14.5	457	1	MUC_SUNMU	P20768	suncus muri	

Biochemistry 9:3188-3196(1970).
 [7]
 RN DISULFIDE BONDS.
 RX MEDLINE=77070267; PubMed=1002129;
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT cyanogen bromide cleavage products, and the disulfide bridges.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 [8]
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE=81208100; PubMed=7236608;
 RA Deisenhofer J.;
 RT "Crystallographic refinement and atomic models of a human Fc fragment
 RT and its complex with fragment B of protein A from Staphylococcus
 RT aureus at 2.9- and 2.8-A resolution.";
 RL Biochemistry 20:2361-2370(1981).
 CC -1- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE
 CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)
 CC MARKER & THE GIM (NON-1) MARKERS.
 CC -1- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
 CC 35,116,198,269 & 272.
 CC -1- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
 CC 268-272.
 CC -1- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
 CC RESIDUES 198,267&272.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J00228; AAC82527.1; ALT_INIT.
 DR PTR; A02146; GHU.
 DR PDB; 1FC1; 15-JUL-92.
 DR PDB; 1FC2; 15-JUL-92.
 DR MIN; 147100; -
 DR INTERPRO; IPR000495; -
 DR INTERPRO; IPR003006; -
 DR PFAM; PF00047; Ig; 3.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 3D-structure.
 FT NON_TER 1 1
 FT DOMAIN 1 98
 FT HINGE 1 98
 FT DOMAIN 99 110
 FT CH1. 111 223
 FT DOMAIN 111 223
 FT CH2. 224 330
 FT DOMAIN 224 330
 FT DISULFID 27 83
 FT DISULFID 103 103
 FT DISULFID 109 109
 FT INTERCHAIN (WITH LIGHT CHAIN). 112 112
 FT DISULFID 112 112
 FT DISULFID 144 204
 FT DISULFID 250 308
 FT DISULFID 180 180
 FT CARBOHYD 97 97
 FT VARIANT 239 239
 FT VARIANT 241 241
 FT MOD_RES 330 330
 FT STRAND 123 126
 FT HELIX 130 134
 FT TURN 136 137
 FT STRAND 141 148
 FT STRAND 158 162
 FT TURN 163 164

Query Match 54.1%; Score 1231; DB 1; Length 330;
 Best Local Similarity 92.5%; Pred. No. 2.9e-76;
 Matches 233; Conservative 3; Mismatches 12; Indels 4; Gaps 2;
 QY 159 THDNICSGSE-SPOKCGIDV---TVDKTHTCCPCAPPELLGGPSVFLFPKPKDTLMIS 214
 Db 78 TQTYCNVNHKPSNTKVKKVEPKSCDKTHTCPCAPPELLGGPSVFLFPKPKDTLMIS 137
 QY 215 RTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKTPREEQYNSTYRVSVLTVLHODWL 274
 Db 138 RTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKTPREEQYNSTYRVSVLTVLHODWL 197
 QY 275 NGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYITLPPSRDELTKNQVSLTCLVKGFYP 334
 Db 198 NGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYITLPPSRDELTKNQVSLTCLVKGFYP 257
 QY 335 SDIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFCSCVMEALHN 394
 Db 258 SDIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFCSCVMEALHN 317
 QY 395 HYTKSLSPG 406
 Db 318 HYTKSLSPG 329
 RESULT 2
 GC2_HUMAN
 ID GC2_HUMAN STANDARD; PRT; 326 AA.
 AC P01859;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-2 CHAIN C-REGION.
 GN IGHG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=82197621; PubMed=6804948;
 RA Ellison J.W., Hood L.E.;
 RT "Linkage and sequence homology of two human immunoglobulin gamma
 RT heavy chain constant region genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
 [2]


```

DR INTERPRO: IPR003006; -
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 50.1%; Score 1138.5; DB 1; Length 327;
Best Local Similarity 71.1%; Pred. No. 4.8e-70;
Matches 224; Conservative 16; Mismatches 30; Indels 45; Gaps 5;

QY 137 APCRKHN-----CSV-----FGLLLQKGNATHDNCISG 166
DB 12 APCRSSTSESTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVQLSSGLYSLSSVTV 71
QY 167 NSES--TOKCGID-----VTVDKT-----HTCPCPAPELGGPSVFLFPKPKDTL 211
DB 72 PSSSLGTYTCNVYDHKPSNTKVDKRVESKYGPPCPAPPEFUGGVSFLFPKPKDTL 131
QY 212 MISRTPEVTCVVVDVSHEDPEVKFNWYGVGVEVHNATKPREQYNSTYRVVSVLTVHLQ 271
DB 132 MISRTPEVTCVVVDVSHEDPEVKFNWYGVGVEVHNATKPREQYNSTYRVVSVLTVHLQ 191
QY 272 DWLNGKEYKCKVSKNALKPAIEKTIISKAKGPQEPQVYTLPPSDELTKNQVSTCLVKG 331
DB 192 DWLNGKEYKCKVSKNALKPAIEKTIISKAKGPQEPQVYTLPPSDELTKNQVSTCLVKG 251
QY 332 FYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEA 391
DB 252 FYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEA 311
QY 392 LHNHYTKSLSPG 406
DB 312 LHNHYTKSLSLG 326

RESULT 4
GC3_HUMAN STANDARD; PRT; 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).
GN IGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RP SEQUENCE (DISEASE PROTEIN WIS).
RX MEDLINE-81021548; PubMed-6774747;
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
RL Biochemistry 19:4304-4308(1980).
[2]
RN REVISIONS TO 12-97 OF PROTEIN WIS.
RX MEDLINE-77118561; PubMed-402363;
RA Michaelson T.E., Frangione B., Franklin E.C.;
RT "Primary structure of the 'hinge' region of human IgG3. Probable
RL J. Biol. Chem. 252:883-889(1977).
[3]
RN REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).
RX MEDLINE-77021516; PubMed-823945;
RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
RT Structure of the FC fragment of immunoglobulin G3.";
RL Biochem. Biophys. Res. Commun. 71:907-914(1976).
[4]
RN SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
RX MEDLINE-82247835; PubMed-6808505;
RA Alexander A., Steinmetz M., Barritault D., Frangione B.,
RA Franklin E.C., Hood L., Buxbaum J.N.;
RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
RL gene deletion model.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
CC NORMALLY PRESENT IN THE HINGE REGION.
CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
CC REF. 2.
CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
CC AND ALL OF THE CH1 REGION.
CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
CC GAMMA-3 HEAVY CHAINS.
CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
CC OR ANOTHER GAMMA CHAIN SUBCLASS.
CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
CC SEGMENT (12-28).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; J00231; AAA52805.1; ALT_SEQ.
PIR; A02149; G3HWTI.
MIM; 147120;
INTERPRO: IPR000495;
INTERPRO: IPR003006;
PFAM; PF00047; Ig; 2.
PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT DOMAIN 12 73 HINGE.
FT DOMAIN 74 183 CH2.
FT DOMAIN 184 289 CH3.
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1
FT CARBOHYD 6 6
FT DISULFID 7 7 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 24 24 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 27 27 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 33 33 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. ).
FT MOD_RES 290 290 REMOVED POST-TRANSLATIONALLY.
FT VARIANT 126 127 QV -> EB (IN ZUC).
FT /FTID-VAR_003890.

```

```
FT VARIANT 134 134 P -> L (IN OMM).
FT FTID-VAR_003891.
FT VARIANT 139 139 F -> Y (IN OMM).
FT FTID-VAR_003892.
FT VARIANT 182 182 T -> A (IN OMM).
FT FTID-VAR_003893.
FT VARIANT 227 227 S -> N (IN OMM).
FT FTID-VAR_003894.
FT VARIANT 227 227 MISSING (IN ZUC).
FT FTID-VAR_003895.
FT VARIANT 279 279 F -> Y (IN OMM).
FT FTID-VAR_003896.
SQ SEQUENCE 290 AA; 32331 MW; E69CBC95705B2F46 CRC64;

Query Match 49.9%; Score 1134.5; DB 1; Length 290;
Best Local Similarity 69.0%; Pred. No. 7.8e-70;
Matches 216; Conservative 19; Mismatches 25; Indels 53; Gaps 4;

QY 97 CLKHRSCTPPPGFGVVOAGTPTERTVCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLT 153
DB 27 CPEPKSCDTPPP-----CPRCPE---PKSCDTPPPC-----54
QY 154 QXGNATHDNCSESTQKGDIVTDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMI 213
DB 55 -----PRCPEPKSCDTPPPCPCPCAPPELLGGPSVFLFPPKPKDTLMI 96
QY 214 SRTPETCVVDVSHEDPEVKENWYDGVVHNNAKTPREEQVNSTYRVSVLTVLHQQW 273
DB 97 SRTPETCVVDVSHEDPEVQFVYDGVVHNNAKTPREEQVNSTYRVSVLTVLHQQW 156
QY 274 LNKKEYCKVSNKALPAPIEKTISKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFY 333
DB 157 LDKKEYCKVSNKALPAPIEKTISKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFY 216
QY 334 PSDIAVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALH 393
DB 217 PSDIAVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALH 276
QY 394 NHYTKSLSLSPG 406
DB 277 NRYTKSLSLSPG 289

RESULT 5
GC_RABIT STANDARD; PRT; 323 AA.
ID GC_RABIT
AC F01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA CHAIN C REGION
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-84030930; PubMed-6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplotype."
RL Immunogenetics 18:387-397(1983).
RN [2]
RN SEQUENCE OF 1-128.
RX MEDLINE-76135469; PubMed-1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit immunoglobulin G of different allotype."
RL Biochem. J. 151:337-349(1975).
RN [3]
RN SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE-83299917; PubMed-6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
```

```
RT RT Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma heavy chain and identification of two genomic C gamma genes; Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RX SEQUENCE OF 132-161.
RX MEDLINE-70110015; PubMed-5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin G."
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (in) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell, Stockholm (1967).
CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER, 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15 MARKERS AND REF.5 THE E15 MARKER.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; M16426; AAA31289.1; -.
DR PIR; A02161; GHRB.
DR INTERPRO; IPR000495; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> E (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).
FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 69B9A118D579A8B CRC64;

Query Match 40.3%; Score 917; DB 1; Length 323;
Best Local Similarity 67.6%; Pred. No. 3.8e-55;
Matches 169; Conservative 31; Mismatches 40; Indels 10; Gaps 3;

QY 167 NSESTQKCGI-----DVTVDKT---HTC---PPCAPPELLGGPSVFLFPPKPKDTLMISRT 216
DB 73 SSSQPVTCNVNHPATNTKVDKTVAPSTCTCPPELLGGPSVFLFPPKPKDTLMISRT 132
QY 217 PEVTCVVVDVSHEDPEVKFNWYDGVVHNNAKTPREEQVNSTYRVSVLTVLHQQW 276
DB 133 PEVTCVVVDVSDQDPEVQFTWYINNEQVTRAPRLPREEQVNSTYRVSVLTPLTHQDWLRG 192
QY 277 KEYCKVSNKALPAPIEKTISKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFVPSD 336
DB 193 KEFKCKVHNKALPAPIEKTISKAGQPLEPKVYTMGPREEELSRVSLTCLMNGYFVPSD 252
QY 337 IAVEWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALH 396
DB 137 IAVEWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALH 396
```


[illegible]

```
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 36.7%; Score 834; DB 1; Length 333;
Best Local Similarity 55.6%; Pred. No. 1.9e-49;
Matches 160; Conservative 41; Mismatches 55; Indels 32; Gaps 5;

QY 130 SNETSSKAPCRKTHNCSVFGLLLTKGNATHDNCSNSESSTCGIDVAVDKTH-----184
DB 66 SSVTSSTWPSQTVT-C-----NVAHP-----ASTK---VDKVERNGGIGH 104

QY 185 -----TCPPCAPPELLGSPSVFLFPKPKDLMISRTPEVTCVVDVSHEDPEVKFNWY 238
DB 105 KCPTCTCHKCPVPELLGSPSVFLFPKPKDILLISQNAKVTVCVVDVSEEDPDVQSF 164

QY 239 VDGVEVINAKTKPREQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISK 298
DB 165 VNNVEVHTAQTPREQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPSPIEKTISK 224

QY 299 AKGQPREQYVTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPPENNYKTTTPVL 358
DB 225 PKGLVRAPQVVMGPPPEQTEQVTSVLTCLTSGELPNDIGVETWSNGHIEKNKNTPEVM 284

QY 359 DSDGSFFLYSLKTVDKSRWQGNVFCVSMHEALHNHYTKLSLSLSPG 406
DB 285 DSDGSFFMYSKLVNRSRWDSPRAFCVSVVHGLNHHVKEKSISSRPPG 332

RESULT 10
GC1_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1998).
DR PIR; PS0017; PS0017.
DR INTERPRO; IPR000495; -.
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 35.9%; Score 817.5; DB 1; Length 326;
Best Local Similarity 60.5%; Pred. No. 1.9e-48;
Matches 144; Conservative 42; Mismatches 37; Indels 15; Gaps 3;

QY 172 QKCGIDVTVDKTHTCPCPAPELLGG---PSVFLFPKPKDLMISRTPEVTCVVDVSH 228
```



```

DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00795; CAA24176.1; -
DR PIR; A02159; GMS.
DR MGD; MGI:96446; IGH-4.
DR INTERPRO; IPR000495; -
DR PFAM; PF00047; I9; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 110
FT DOMAIN 111 217
FT DOMAIN 218 324
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT MOD_RES 324 324
FT CONFLICT 276 276
FT CONFLICT 278 278
SQ SEQUENCE 324 AA; 35704 MW; A38812F3D1F2C93 CRC64;

Query Match 35.7%; Score 812.5; DB 1; Length 324;
Best Local Similarity 62.3%; Pred. No. 4.2e-48;
Matches 139; Conservative 44; Mismatches 35; Indels 5; Gaps 2;

QY 186 CPP--CPAPELGGPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNMYDQVE 243
DB 104 CKPCICTVPEV---SSVFIPPKPKDVLITLTPKVTCCVVVDISKDDPEVQFSWFVDVDE 160

QY 244 VNAKTKPREQNSYRVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISKAKGP 303
DB 161 VHTAQTPREEQNSFRSSELPIMHQDLNGKEFKCRVNSAFAPIETKISKTKGRP 220

QY 304 REPQVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPTPPVLDSGDS 363
DB 221 KAQVTTIPPKQMAKDKVSLTCHMTDFPEPDTVWQNGQPAENYKNTQPLIMNTGS 280

QY 364 FFYLSKLTVDKSRWQGNVFCSSVMHEALNHYTKRSLSPG 406
DB 281 YFVYSKLVNOKSNWEAGNTFTCSVLHGLNHHNTERKSLSPG 323

RESULT 12
GCIM_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-80045036; PubMed-115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568 (1979).
RN [2]
RN SEQUENCE OF 323-393 FROM N.A.

```

```

RX MEDLINE-82197626; PubMed-6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RN SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE-82115295; PubMed-6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RN SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE-82222190; PubMed-6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -I- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
SEGMENT OF MU CHAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V00793; CAA24172.1; -
CC EMBL; V00793; CAA24173.1; -
CC EMBL; V00793; CAA24174.1; -
CC PIR; B02159; GMSM.
CC MGD; MGI:96446; IGH-4.
CC INTERPRO; IPR000495; -
CC INTERPRO; IPR003006; -
CC PFAM; PF00047; I9; 3.
CC PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 110
FT DOMAIN 111 217
FT DOMAIN 218 324
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 35.7%; Score 812.5; DB 1; Length 393;
Best Local Similarity 62.3%; Pred. No. 5.1e-48;
Matches 139; Conservative 44; Mismatches 35; Indels 5; Gaps 2;

QY 186 CPP--CPAPELGGPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNMYDQVE 243
DB 104 CKPCICTVPEV---SSVFIPPKPKDVLITLTPKVTCCVVVDISKDDPEVQFSWFVDVDE 160

```

QY 244 VHNKTRPREQYNYSTRVSVLVLHODWLNKGYCKYVSNKALPAPIEKTISKAKGP 303
DB 161 VHTAQTPREQNFSTERSVSELPIMHQDWLNKGYCKYVSNKALPAPIEKTISKGRP 220
QY 304 REPQYVTLPPSRDELTKNOVSLTCLVKGFFPSDITAVEWESNGQPNYKTPPVLDSDGS 363
DB 221 KAPOYVITIPPKEQMAKDKVSLTCMIDFFPEDITVEMQWNGQPAENYKTPQIMNTNGS 280
QY 364 FFLYSLKTVKDSRWQOQVFCVSNVHEALHNHYTKOKLSLSLSPG 406
DB 281 YFVYSKLVNQSWEAGNFTFCSVLHGLHNHTEKLSLSHSPG 323

RESULT 13
GCC_RAT
ID GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2C CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RL Eur. J. Immunol. 18:317-319(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X07189; CAA30169.1; -
DR PIR: S00847; S00847.
DR INTERPRO: IPR000495; -
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 35.4%; Score 805; DB 1; Length 329;
Best Local Similarity 57.7%; Pred. No. 1.4e-47;
Matches 146; Conservative 44; Mismatches 49; Indels 14; Gaps 2;
QY 168 SESTQKGDIVYDKTHTC-----pp---CPAPELLGGPSVFLPPKPKDILMI 213
DB 76 SSQVTCVSAHPAYKSNLIKRIEPRPKPRPTDTCSCDNDNLGRPSVFIFPPKPKDILMI 135
QY 214 SRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNAKTPREQYNSYTRVSVLTVLHODW 273
DB 136 TLTKPVTCVVVDVSEEDPDVQSFVFNVRVFTAQTPHEQLNGTRFVSTLHIQHDW 195

QY 274 LNKKEYCKYVSNKALPAPIEKTISKAKGPQVYTLPPSRDELTKNOVSLTCLVKGFFY 333
DB 196 MSGREFKCKVNNKDLPSPIEKTISKPRKARTPOVYITIPPPREQMSKNKVSILTQWTSFY 255
QY 334 PSDITAVEWESNGQPNYKTPPVLDSDGSFFLYSKLTVKDSRWQOQVFCVSNVHEALH 393
DB 256 PASISVEMERNGELEDYKNTLPVLDSDESYFLYSLKSLSDVTDSWMRGDIYTCVSVHEALH 315
QY 394 NHYTKOKLSLSLSPG 406
DB 316 NHHTQKNLSRSPG 328

RESULT 14
GCAA_MOUSE
ID GCAA_MOUSE STANDARD; PRT; 330 AA.
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Nucleic Acids Res. 8:3143-3155(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
RT and evolution of heavy chain genes: further evidence for intervening
RT sequence-mediated domain transfer.";
RL Nucleic Acids Res. 9:1365-1381(1981).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=8123894; PubMed=6787604;
RA Ollio R., Auffray C., Mochamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
RT suggests that exons can be exchanged between genes in a multigenic
RL family.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN [4]
RP MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin: amino-acid sequence of the FC fragment. Implications
RL for the evolution of immunoglobulin structure and function.";
RN Eur. J. Biochem. 43:423-435(1974).
RN [5]
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
RT immunoglobulin. Identification of the disulfide bridges.";
RN Eur. J. Biochem. 30:452-462(1972).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: V00798; CAA24178.1; -
DR PIR: A02152; G2MSA.

```
DR INTERPRO: IPR000495; -
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;
```

```
Query Match 35.4%; Score 804; DB 1; Length 330;
Best Local Similarity 65.0%; Pred. No. 1.6e-47;
Matches 145; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

QY 186 CPP--CPAPELLGGPSVLEFPKPKDTLMTSRPEVTCVVVDVSHEDPEVKFNMYVDGVE 243
Db 107 CPPCKCPAPNLLGGPSVFIPPKIKDVLMTSLSPITVCVVVDVSEDDPDVQISWFVNVE 166

QY 244 VHNAKTKPREEQNSTYRVSVLTVLHQLNGKEYCKVSKNKPAPIEKTIKSKAGQP 303
Db 167 VHTAQQTQTHREDYNSLTVVVSALPIQHDWMSGKEFKCKVKNKDLPAPIERTISKPKGSV 226

QY 304 REPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDAVAVESNGOPENNYKTPPVLDSDGS 363
Db 227 RAPQVYVLPPEEMTKKQVLTLCMTDFMPEDIIYVETWNGKTELNYKNTPEVLDSGS 286

QY 364 FFYLSKLTVDKSRWQGNVFCSCVMHEALNHYTKQSLSPG 406
Db 287 YFMYSKLRVEKKNWERNYSYSCSVVHEGLNHHHTTKFSRTPG 329
```

```
RESULT 15
GCAM_MOUSE STANDARD; PRT; 399 AA.
AC F01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-8222190; PubMed-6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MENA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC ----- J00471; AAB59661.1; ALT_INIT.
DR EMBL; J00471; AAB59661.1; G2MSAM.
DR PIR; A02154; G2MSAM.
DR MGD; MGI:96443; IGH-1.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 364 399
FT DOMAIN 180 180
FT CARBOHYD 180 180
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAD3FF0 CRC64;

Query Match 35.4%; Score 804; DB 1; Length 399;
Best Local Similarity 65.0%; Pred. No. 1.9e-47;
Matches 145; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

QY 186 CPP--CPAPELLGGPSVLEFPKPKDTLMTSRPEVTCVVVDVSHEDPEVKFNMYVDGVE 243
Db 107 CPPCKCPAPNLLGGPSVFIPPKIKDVLMTSLSPITVCVVVDVSEDDPDVQISWFVNVE 166

QY 244 VHNAKTKPREEQNSTYRVSVLTVLHQLNGKEYCKVSKNKPAPIEKTIKSKAGQP 303
Db 167 VHTAQQTQTHREDYNSLTVVVSALPIQHDWMSGKEFKCKVKNKDLPAPIERTISKPKGSV 226

QY 304 REPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDAVAVESNGOPENNYKTPPVLDSDGS 363
Db 227 RAPQVYVLPPEEMTKKQVLTLCMTDFMPEDIIYVETWNGKTELNYKNTPEVLDSGS 286

QY 364 FFYLSKLTVDKSRWQGNVFCSCVMHEALNHYTKQSLSPG 406
Db 287 YFMYSKLRVEKKNWERNYSYSCSVVHEGLNHHHTTKFSRTPG 329
```

Search completed: March 1, 2001, 09:17:47
Job time: 265 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:46 ; Search time 79.26 seconds
(without alignments)
347.813 Million cell updates/sec

Title: US-09-389-782A-6
Perfect score: 2274
Sequence: 1 ETPPKYLHYDEETSHOLL.....VMHEALHHYQKSLSPG 406

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231	54.1	330	1 GHU	Ig gamma-1 chain C
2	1228	54.0	255	4 S31866	Ig gamma-1 chain C
3	1227	54.0	374	2 S72664	Ig heavy chain V r
4	1222.5	53.8	374	2 S69339	Ig heavy chain V r
5	1180	51.9	234	2 PT0207	Ig gamma chain C r
6	1156	50.8	377	2 A23511	Ig gamma-3 chain C
7	1154	50.7	377	2 A60764	Ig gamma-3 chain C
8	1148	50.5	326	1 G2HU	Ig gamma-2 chain C
9	1138.5	50.1	327	1 G4HU	Ig gamma-4 chain C
10	1134.5	49.9	289	1 G3HUW1	Ig gamma-2 chain C
11	917	40.3	323	1 GHRB	Ig gamma-3 heavy c
12	907	39.9	328	2 I47160	Ig gamma chain C r
13	907	39.9	328	2 I47159	Ig gamma-2b chain
14	905	39.8	277	2 I47162	Ig gamma 2a chain
15	895.5	39.4	329	1 G2GP	Ig gamma 4 chain c
16	891.5	39.2	328	2 I47158	Ig gamma-2 chain c
17	879	38.7	328	2 I47161	Ig gamma 1 chain c
18	858.5	37.8	470	2 S22080	Ig gamma 3 chain c
19	848.5	37.3	329	1 G3MSC	Ig heavy chain pre
20	842.5	37.0	398	1 G3MSM	Ig gamma-3 chain C
21	841	37.0	308	2 C30554	Ig gamma-3 chain C
22	841	37.0	472	2 S31459	Ig heavy chain C r
23	834	36.7	333	2 PC0018	Ig gamma-1 chain -
24	822.5	36.2	444	2 PC4436	Ig gamma-2b chain
25	817.5	35.9	326	2 PS0017	monoclonal antibod
26	812.5	35.7	324	1 G1MS	Ig gamma-1 chain C
27	812.5	35.7	393	1 G1MS	Ig gamma-1 chain C
28	805	35.4	329	2 S00847	Ig gamma-2c chain
29	804	35.4	330	1 G2MSA	Ig gamma-2a chain

ALIGNMENTS

RESULT 1
GHU

Ig gamma-1 chain C region - human
C:Species: Homo sapiens (man)

C>Date: 31-Jan-1981 #sequence revision 18-Aug-1982 #text change 16-Jul-1999
C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146
R:Ellison, J.W.; Beison, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:217370

A:Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) marker
A:Note: Lys-330 is removed after translation

R:Harris, L.J.

submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <HAR>

A:Cross-references: EMBL:217370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of

A:Reference number: S33887; MUID:83001943

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113;235-330 <TAK>

A:Cross-references: EMBL:217370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq

A:Reference number: A90563; MUID:71064024

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96,'R',98-135 <CUN>

A:Note: this sequence has the G1m(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se

A:Reference number: A90564; MUID:71064025

A:Contents: Eu

A:Accession: A90564

A:Molecule type: protein

A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2

A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met

R:Ponstingl, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A:Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni

Query Match	40.3%	Score 917;	DB 1;	Length 323;
Best Local Similarity	67.6%;	Pred. No. 1.5e-49;		
Matches 169;	Conservative 31;	Mismatches 40;	Indels 10;	Gaps 3;
QY	167	NSESTQKCGI-----DVTVDK	---HTC--PPCPAPELLGGPSVFLPPPKPKDILMISRT	216
	:	:	:	:
Db	73	SSSPVTCNVAHPATNKVDTPASTCSKPTCPPELLGGPSVFIFPPKPKDILMISRT	132	
QY	217	PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVYSVLTVTLQHDWLG	276	
	:	:	:	:
Db	133	PEVTCVVVDVSDQDEVOFTWYINNEQVITARPPLRQQQNFSTIRVYSTLPIITHQDMLRG	192	
QY	277	KEYKCKYSNKAIPAEIKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVVGFYPSD	336	
	:	:	:	:
Db	193	KEFKCKYHNKALPAIEIKTISKAKGQPLEPKVYITMGPPREELSSRSVSLTCWINGFYPSD	252	
QY	337	IAVESWGSPENNYKTPPVLSDSGSFYFLYKLTVDKSRMQQGNVFVSCSVMHREALHNHY	396	
	:	:	:	:
Db	253	ISVEMWNGRAEDNYKTPPAVLDSGGYFLYKLSVPTSEWQRGDVTCVSMHEALHNHY	312	
QY	397	TQKSLSLSPG	406	
	:	:	:	:
Db	313	TQKISIRSPG	322	

Query Match	39.9%;	Score 907;	DB 2;	Length 328;
Best Local Similarity	65.9%;	Pred. No. 6.1e-49;		
Matches 170; Conservative	33;	Mismatches 35;	Indels 20;	Gaps
Qy 165	SGNSESTQKGID-----VTVDK---THTCPPCP-----	APELLGGPSVFLPFRPKDT	210	
Db 74	SSLSSKSYTCNVNHPATTTKVDKRVGKTKPPCIPACESP---	GPSVFIFPRPKDT	129	
Qy 211	LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP	PREEQYNSTKYRVVSVLT	270	
Db 130	LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP	PREEQYNSTKYRVVSVLT	189	
Qy 271	QDWLNGKEYECKVSNKALPAPIEKTISKAKGQPREQVYTL	PSRDELTKNOVISLFLCKV	330	
Db 190	QDWLNGKEYECKVSNKALPAPIEKTISKAKGQPREQVYTL	PSRDELTKNOVISLFLCKV	249	
Qy 331	GFYPSPDITAVESWNGQ--PENNNKTTTPPVLLDSGSF	FLYSKLTVDKSRWQOQNVFCSVM	368	
Db 250	GFYPSPDITAVESWNGQ--PENNNKTTTPPVLLDSGSF	FLYSKLTVDKSRWQOQNVFCSVM	309	
Qy 389	HEALHNHYTQKSLSLSPG	406		
Db 310	HEALHNHYTQKSLSLSPG	327		

RESULT 13
I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:KacsKovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845
A:Accession: I47159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:A/Cross-references: EMBL:U03779; NID:9433123; PTDN:AAA52217.1; PID:9433124
C:Genetics:
C:Gene: IgG2a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IIM>

Query Match	39.9%;	Score 907;	DB 2;	Length 328;
Best Local Similarity	65.9%;	Pred. NO. 6.1e-49;		
Matches 170; Conservative	33;	Mismatches 35;	Indels 20;	Gaps 5;

Db 221 KGAPRMPDVYTLPPSRDELSKSVSVTCIIINFFPADIHVWASNRVPVSEKEYKNTPEPI 280
Qy 358 LDSGSGFFLYSKLTVDKSRWQQGNVFCVMHEALHNHYTOKSLSLSPG 406
Db 281 EDADGSYFLYSKLTVDKSAWDQGTYYTCSVMHEALHNHYTOKAISRSFG 329

Search completed: March 1, 2001, 09:15:48
Job time: 146 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:03 ; Search time 70.93 Seconds
(without alignments)
195.724 Million cell updates/sec

Title: US-09-389-782A-6
Perfect score: 2274
Sequence: 1 ETTPPKYLHYDEETSHQLC.....VMHEALHHVTKSLSPG 406

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1490	65.5	485	13	R24016
2	1487	65.4	518	15	R51003
3	1349	59.3	376	19	W60037
4	1334	58.7	376	18	W30287
5	1330	58.5	438	16	R81882
6	1330	58.5	438	19	W48976
7	1283	56.4	764	21	Y68949
8	1257.5	55.3	664	19	W71603
9	1257.5	55.3	664	21	Y80123
10	1254.5	55.2	664	19	W71604
11	1254.5	55.2	664	21	Y80124
12	1253	55.1	963	19	W70540

13	1250.5	55.0	388	19	W73513
14	1250.5	55.0	388	21	Y54063
15	1250	55.0	592	20	W70797
16	1250	55.0	592	21	Y92185
17	1247	54.8	535	20	Y17414
18	1246.5	54.8	680	15	R48037
19	1246	54.8	445	20	Y24153
20	1244	54.7	388	19	W73514
21	1244	54.7	388	21	Y54064
22	1241	54.6	347	13	R27163
23	1241	54.6	347	21	Y83136
24	1240.5	54.6	449	20	W96278
25	1239.5	54.5	784	21	Y92207
26	1238.5	54.5	375	18	W06683
27	1238.5	54.5	408	21	Y44461
28	1238	54.4	595	20	W86003
29	1236.5	54.4	387	17	R90920
30	1235	54.3	396	18	W18574
31	1235	54.3	396	18	W18575
32	1234	54.3	399	21	Y70867
33	1234	54.3	400	21	Y15123
34	1234	54.3	446	15	R58753
35	1234	54.3	446	20	Y23986
36	1234	54.3	446	20	Y01037
37	1234	54.3	446	20	W96743
38	1233	54.2	371	10	P93558
39	1233	54.2	423	21	Y70869
40	1233	54.2	460	21	Y69890
41	1233	54.2	480	20	W90206
42	1232.5	54.2	482	19	W31646
43	1232	54.2	235	20	Y01372
44	1232	54.2	418	21	Y70868
45	1231	54.1	329	17	R91806

ALIGNMENTS

RESULT 1	
R24016	
ID R24016 standard; Protein; 485 AA.	
XX	
AC R24016;	
XX	
DT 26-NOV-1992 (first entry)	
XX	
DE Fusion protein TNFRFc.	
XX	
KW Tumour necrosis factor; TNF; IgG1; immunoglobulin G1.	
XX	
OS Synthetic.	
XX	
FH Key	Location/Qualifier
FT Region	1..252
FT	/note= "human TNF extracellular receptor"
FT Region	253..270
FT	/note= "linker and hinge"
FT Region	271..379
FT	/note= "IgG1 CH2"
FT Region	380..485
FT	/note= "IgG1 CH3"
XX	
PN EP488170-A.	
XX	
PD 03-JUN-1992.	
XX	
PF 26-NOV-1991;	91EP-0120187.
XX	
PR 28-NOV-1990;	90DE-4037837.
XX	
PA (BEHW) BEHRINGWERKE AG.	
XX	
PI Lauffer L, Oquendo P, Zettlmeissl G;	


```

Db 368 nstyrvvsvltvlnqdwngdyckvsnkalpampgktiskakggpreqvyltppsrd 427
QY 317 ELTKNOVSLTCLVKGFPSPIAWEWESGOPENNYKTPPVLDSDGSFFLYSLKLTVDKSR 376
Db 428 eltknqvsllclvkgyfprhriavewesngopennykttppvldsdgsfflyslkltvdksr 487
QY 377 WQGNVFCSCVMHEALHNHYTKQSLSLSPG 406
Db 488 wqggnvfscsvmhealhnhytqkslsisp 517

RESULT 3
W60037 ID W60037 standard; Protein; 376 AA.
XX AC W60037;
XX DT 11-SEP-1998 (first entry).
XX DE Antigenic peptide hFas (nd29) containing Fc region.
XX KW Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
XX KW heart failure; kidney failure; graft-versus-host disease; antibody;
XX KW myocardial infarction; ischemic restenosis; endotoxemic shock.
XX OS Homo sapiens.
XX FH Key
XX FT Peptide
XX FT /note= "hFas antigen signal peptide"
XX FT 30..376
XX FT /note= "hFas (nd29) protein"
XX PN W60037
XX PD 07-MAY-1998.
XX PF 31-OCT-1997; 97WO-JP03978.
XX PR 26-SEP-1997; 97JP-0262521.
XX PR 31-OCT-1996; 96JP-0290459.
XX PR 27-DEC-1996; 96JP-0351718.
XX PA (MOCH ) MOCHIDA PHARM CO LTD.
XX PA (OSAB-) OSAKA BIOSCIENCE INST.
XX PI Nagata S, Suda T, Yatomi T;
XX WPI; 1998-271925/24.
XX DR N-PSDB; V34430.
XX PT Use of Fas antagonist for treatment and prevention of
XX PT apoptosis-related diseases - such as heart or kidney failure,
XX PT graft-versus-host disease or liver disease
XX PS Examples; Fig 5-9; 86pp; Japanese.
XX CC This represents the antigenic peptide hFas (nd29) containing the Fc
XX CC region. The invention provides the use of Fas antagonist as an agent for
XX CC the treatment and prevention of apoptosis-related diseases. The Fas
XX CC antagonist can be a partial Fas antigen peptide containing the
XX CC extracellular part of the protein, but lacking the signal sequence, an
XX CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is
XX CC preferably a humanized antibody. The Fas antagonist is used in the
XX CC treatment and prevention of diseases such as myocardial infarction, heart
XX CC failure, ischemic heart disease, acute kidney failure, graft-versus-host
XX CC disease, ischemic restenosis of the heart, liver or kidney, and
XX CC endotoxemic shock, and also as an organ preservative in transplantation.
XX CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand
XX CC system.
XX SQ Sequence 376 AA;

```

```

Query Match 59.3%; Score 1349; DB 19; Length 376;
Best Local Similarity 66.3%; Pred. No. 1.5e-76;
Matches 270; Conservative 15; Mismatches 60; Indels 62; Gaps 9;
QY 8 LHYDEETSHQLLCKCPPTGLYKQHTAKW-KTVCAPCPD-HYYTDSMHTSDEGLYCSVP 65
Db 23 lhdgqfchk----pcppgerkardctvngdepdcvpcqegkeytdkafsskrrcr-l 77
QY 66 CKELQYVYKQE--CNRTNRRVCECKEGRYLEIEFCLKHRSCPP-----GFGVVQAGTPPERT 119
Db 78 cdegghgleivnctrtqtkcrckpnffonstvc---ehcdptckcehglikeetltsnt 134
QY 120 VCKRCPDGFSSNETSSKAPCRKHTNCVSFGLLLTKGNATHDNCISGNESTQKCGIDVT 179
Db 135 kcke--egrsnepks----- 148
QY 180 VDKTHTCPCPAPPELLGGPSVFLPPPKPDTLMISRTPEVTVVVDVSHEDPEVKFNMYV 239
Db 149 cdkthtccppcappelggpsvflfppkpkdtlmisrtpevtcvvvdvshedpevfwnyvv 208
QY 240 DGEVHNNAKTPREEQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETISK 299
Db 209 dgvevhnaktpreegynstyrvvsvltvlhqdwlngkeyckvsnkalpapietiska 268
QY 300 KGQPREPQVYTLPSRDELTKNQVSLTCLVKGFYPSDIAWESNGOPENNYKTTTPPVLD 359
Db 269 kgqprepqvylppsrdeltknqvsllclvkgyfypsdiawesngopennykttppvld 328
QY 360 SDGSFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPG 406
Db 329 sdgsfflyslkltvdksrwqggnvfscsvmhealhnhytqkslsisp 375

RESULT 4
W50287 ID W50287 standard; Protein; 376 AA.
XX AC W50287;
XX DT 16-JUL-1998 (first entry)
XX DE Human Fas antigen derivative/IgG1 Fc fusion.
XX KW Human; Fas antigen; derivative; apoptosis regulation; gene therapy;
XX KW treatment; diabetes; arthritis; lupus; hepatitis; influenza; HIV;
XX KW apoptosis modulation; immunoglobulin G1 Fc; IgG1 Fc; fusion.
XX OS Homo sapiens.
XX FH Key
XX FT Peptide
XX FT /label= sig_peptide
XX FT 17..376
XX FT /label= mat_peptide
XX PN W09742319-AL.
XX PD 13-NOV-1997.
XX PF 01-MAY-1997; 97WO-JP01502.
XX PR 02-MAY-1996; 96JP-0135760.
XX PA (MOCH ) MOCHIDA PHARM CO LTD.
XX PA (OSAB-) OSAKA BIOSCIENCE INST.
XX PI Nagata S, Nakamura N;
XX WPI; 1997-558981/51.
XX DR N-PSDB; V07004.
XX

```

PT Fas antigen derivative containing modified extracellular region -
PT has low antigenicity, promotes apoptosis and is useful in treatment
of viral and other diseases
XX
PS Disclosure; Fig 4; 102pp; Japanese.
XX
CC The present sequence is a Fas antigen derivative/IgG1 Fc
CC fusion, which contains a Fas antigen extracellular region lacking
CC one or more amino acid residues in the region from the
CC amino-terminal to (but excluding) the 1st cysteine residue
CC (preferably at least 29 residues are deleted).
CC The derivative is an effective regulator of apoptosis and can be
CC used (either by administration of the polypeptide, or by the use
CC of the coding DNA in gene therapy) to treat a range of diseases,
CC e.g. diabetes, arthritis, lupus and in particular viral diseases
CC such as hepatitis, influenza and HIV, by modulating apoptosis of
CC virus-infected cells.
XX
SQ Sequence 376 AA;

Query Match 58.7%; Score 1334; DB 18; Length 376;
Best Local Similarity 62.6%; Pred. No. 1.3e-75;
Matches 267; Conservative 16; Mismatches 62; Indels 62; Gaps 9;

Qy 8 LHYDEETSHQLLCKDPCPGTYLKHCTAKW-KTVCAPCPD-HYITDSWHTSDECLYCSVPV 65
Db 23 lhhgqfchk-----pcppgkardctvngdepvcpcqegkeytdkafskrrcr-i 77
Qy 66 CKEIYVQKE--CNRTHNRVCEKGRYLEIEFCLKHSRCP-----GFGVVQAGTPERT 119
Db 78 cdegghleincrtqtkrcpfnfstvc---ehcdpctkcehglikectltnt 134
Qy 120 VKRCPCDFFSNESKAPCRKHTNCVFGILLTLQKGNATHDNCSGSESTQKCGIDVT 179
Db 135 kcke--egrsnepks----- 148
Qy 180 VDKHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYV 239
Db 149 cdktctcpapellgspvsflfppkpkdtlmsptpevtcvvdvshedpevkfnmyv 208
Qy 240 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIKA 299
Db 209 dgvevhnaktprreeqynstyrvvsvltvlhqdwingkeyckvsnkalpapiektiska 268
Qy 300 KGPQREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 359
Db 269 kgpqrepqytlppsrdeitknqvsiltclvkgyfypsdiavewesngqpennnykttppvld 328
Qy 360 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYTKQSLSLSPG 406
Db 329 sdgsfflyslntvdksrwqqgnvfscsvmhealhhnytkqslslspg 375

RESULT 5
R81882
ID R81882 standard; Protein; 438 AA.
XX
AC R81882;
XX
XX
DT 30-MAR-1996 (first entry)
XX
DE Plasmid pDC406/OX40/Fc* encoding an OX40/Fc mutein protein.
XX
KW OX40; OX40-L; cytokine; cell surface molecule; plasmid;
KW pDC406/OX40/Fc*; membrane glycoprotein.
XX
OS Synthetic.
XX
PN US5457035-A.
XX
PD 10-OCT-1995.
XX

PF 23-JUL-1993; 93US-0097827.
PR 23-JUL-1993; 93US-0097827.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
XX
DR WPI: 1995-357992/46.
DR N-PSDB: T00829.
XX
PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
PT and host cells, used to produce recombinant ligand used in e.g.
PT prim. T cell culture, to modulate immune response etc.
XX
PS Example 2; Column 35-38; 26pp; English.
XX
CC This plasmid encodes an OX40/Fc antibody fragment mutein protein,
CC and is used to express a soluble OX40/Fc mutein fusion protein for
CC use in detecting cDNA clones encoding a OX40 ligand. The Fc
CC fragment may be derived from human IgG1, and the plasmid may be
CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell
CC line. Culture supernatant was purified by affinity chromatography
CC and this was used, together with labeled goat anti-human IgG to
CC screen various cell lines.
XX
SQ Sequence 438 AA;

Query Match 58.5%; Score 1330; DB 16; Length 438;
Best Local Similarity 62.6%; Pred. No. 2.7e-75;
Matches 263; Conservative 23; Mismatches 80; Indels 54; Gaps 6;

Qy 20 CDKCPPTGLYKQHTAKWKTCAPCPDHYITD--SWHTSDECLYCSVPCKELQYVQECN 77
Db 39 ccreqpgghgmvrchdtrdtlchpcetgfneavnydtkctqcnh--rsgselkgnct 96
Qy 78 RTHNRVCEKGRYLEIEFCLKHSRCPGFGVQAGTPERTNVCKRCPDGPFPSNETSKA 137
Db 97 ptqdtvcr-----pgtprqdsqykglydcvpcppghfs--pgnnq 137
Qy 138 PCRHTNGSVFGLLLTLQKGNATHDNCSGNS-----ESTQKCGIDVT----- 180
Db 138 ackptwntclsgkqtrhpasdsldavcedrsilatllwetqrptfrptvgttwwpts 197
Qy 181 -----DKHTCTCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDV 226
Db 198 elpattptliveprscdkthtccpapeagapsvflfppkpkdtlmsrtpevtcvvvdv 257
Qy 227 SHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNK 286
Db 258 shedpevkfnmyvdgvevhnaktprreeqynstyrvvsvltvlhqdwingkeyckvsnk 317
Qy 287 ALPAPIEKTIKAKGPQREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 346
Db 318 alpapiektiskakgpqrepqytlppsrdeitknqvsiltclvkgyfypsdiavewesngq 377
Qy 347 PENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYTKQSLSLSPG 406
Db 378 pennnykttppvldsdgsfflyskltvdksrwqqgnvfscsvmhealhhnytkqslslspg 437

RESULT 6
W48976
ID W48976 standard; Protein; 438 AA.
XX
AC W48976;
XX
DT 25-SEP-1998 (first entry)
XX
DE OX40/Fc mutein.
XX
KW OX40/Fc; cytokine; T cell antigen; TH-2 immune response; OX40-L;

XX 19-MAY-2000 (first entry)
 XX Human NTNR alpha and IgG fusion protein SEQ ID NO:16.
 DE Human; neuriturin receptor alpha; NTNR alpha; splenic haematopoiesis;
 KW anaemia; thrombocytopaenia; hypoplasia; haemorrhage.
 XX Homo sapiens.
 OS Synthetic.
 XX US6025157-A.
 XX 15-FEB-2000.
 XX 24-OCT-1997; 97US-0957063.
 PR 18-FEB-1997; 97US-0038839.
 PR 09-JUN-1997; 97US-0049818.
 XX as enhancing splenic haematopoiesis, treating anaemia,
 PA (GETH) GENENTECH INC.
 XX Hynes MA, Rosenthal A, Klein RD;
 XX WPI; 2000-181808/16.
 DR N-PSDB; 291459.
 XX Isolated nucleic acid molecule encodes a neuriturin receptor-alpha amino
 PT acid sequence excluding the N-terminal signal peptide -
 XX Claim 3; Column 91-96; 78pp; English.
 PS The present invention describes a neuriturin receptor alpha (NTNR alpha).
 CC NTNR alpha binds neuriturin. The NTNR alpha nucleic acid molecule is
 CC useful for the expression of NTNR alpha, which may be used to identify
 CC agonist and antagonist compounds having therapeutic applications, such
 CC as enhancing splenic haematopoiesis, treating anaemia,
 CC thrombocytopaenia, hypoplasia, or haemorrhage. The present sequence
 CC is a NTNR alpha and IgG (immunoglobulin G) fusion protein from the
 CC present invention.
 XX Sequence 664 AA;
 SQ

Query Match 55.3%; Score 1257.5; DB 21; Length 664;
 Best Local Similarity 62.7%; Pred. No. 1.3e-70;
 Matches 264; Conservative 25; Mismatches 55; Indels 77; Gaps 14;

QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDRCLYCSPVKRELQV 72
 DB 274 ncrasyqtvtst-cpadnyqacilgsvagmifdmtpnyvdsptg---ivvspwc----- 323
 QY 73 KQECNTHNRVCEKGEGRYLEIEF-----CLKH-----RSCP--PGFGVVOAGTP 115
 DB 324 --scrgsgnmeece--kflr-dftenpcrlnaigafngtdvvnvspkqpsfqatqprv 378
 QY 116 ERNTVCKRCPDFFNETSSKAPCRKHTNCVSFGLLLTOKGNATHNICSNSSTQKCG 175
 DB 379 ext---pslpdd-lsdstslgts-----vittctsvqeqglkannskelsmcf 422
 QY 176 IDVT-----VDKTHCPCPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVD 225
 DB 423 teltniipgrdpvdkthtccpcapellggsfvlfppkpkdtlmisrtpevtcvvvd 482
 QY 226 VSHDEPEVKFNVDGVVHNAKTPREOYNSTYRVSVLTVLHDWLNKGEYKCKVSN 285
 DB 483 vshedpevkfnvvdgvevhnaktkpreeqynstyrsvsvltvlhqdwlngkeyckvsn 542
 QY 286 KALPAPIETKISAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNG 345
 DB 543 kalpapietkiskakgqprepqvtytlppsrdeentknqsltclvkgyfypsdiavewesng 602
 QY 346 QPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFCFSVMHEALHNHYTKLSLSLSP 405

Db 603 qpennyktppvldsdgsfflyskltvdksrwqgnvfcfsvmhealhnhytklsislsp 662
 QY 406 G 406
 Db 663 g 663
 RESULT 10
 W71604
 ID W71604 standard; Protein; 664 AA.
 XX W71604;
 XX 19-NOV-1998 (first entry)
 XX Rat neuriturin receptor alpha/Fc sequence (Iff2a) fusion protein.
 DE Human; neuriturin receptor alpha; NTNR-alpha; variant; chimeric;
 KW fusion protein; immunoadhesion; ret-expressing cell; neurological;
 KW renal; haematological disease.
 XX Synthetic.
 OS Homo sapiens.
 OS Rattus sp.
 XX WO9836072-A1.
 XX 20-AUG-1998.
 XX 17-FEB-1998; 98WO-US03179.
 PR 24-OCT-1997; 97US-0957063.
 PR 18-FEB-1997; 97US-0802805.
 PR 09-JUN-1997; 97US-0871913.
 XX (GETH) GENENTECH INC.
 XX Hynes MA, Klein RD, Rosenthal A;
 XX WPI; 1998-467175/40.
 DR N-PSDB; V58007.
 XX New polypeptide(s) based on human neuriturin receptor alpha and
 PT related nucleic acid - useful for increasing survival of
 PT ret-expressing cells for treating e.g. neurological, renal and
 PT haematological diseases
 XX Example 3; Page 84-86; 116pp; English.
 PS The present sequence represents rat neuriturin receptor alpha
 CC (NTNR-alpha)/Fc sequence (Iff2a) fusion protein, from an example of the
 CC present invention. NTNR-alpha proteins can be used: (a) to identify
 CC molecules that bind specifically to it (potential agonists and
 CC antagonists) and to purify such compounds; (b) to modulate response of
 CC cells to neuriturin (NTN); (c) to increase survival of Ret-expressing
 CC cells or to activate Ret on the surface of cells (soluble glial derived
 CC neurotrophic factor receptor (GDNFR alpha) may be used the same way);
 CC (d) to increase the half-life of cognate ligands (especially NTN); (e)
 CC diagnostically to determine serum levels of its ligands; and (f) as
 CC animal feed additive or molecular weight marker. NTNR-alpha, its genes,
 CC (ant)agonists and antisense nucleic acids, are useful in vivo or ex vivo
 CC for treating conditions related to abnormal NTN activity or response,
 CC particularly neurological (central or enteric), renal or haematopoietic
 CC (spleen) diseases or injuries. Ab may be agonists or antagonists for
 CC therapeutic use (e.g. as antagonists to treat excessive/unwanted NTN
 CC alpha expression, e.g. in some tumours), also reagents for immunoassay
 CC and affinity purification.
 XX Sequence 664 AA;
 SQ

Query Match 55.2%; Score 1254.5; DB 19; Length 664;

Best Local Similarity 62.8%; Pred. No. 2e-70; Matches 262; Conservative 30; Mismatches 56; Indels 69; Gaps 13;

QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDECLYSPVCKELQYV 72
Db 274 ncrasyrtits-cpadnyqacslgyagmifdmtpnvydsnptg---ivspwc----- 323

QY 73 KQECNRTNHRVCECKEGRYLEIEF---CLKHRSPPGFGVQAGTPERNTVCKRCPCDGF 128
Db 324 --ncrgsgnmeeeee--kflr-dftenpcl--rnaiaqfg-----ngtdvnmsspgk- 367

QY 129 FSNETSskapcrkht-----NCSVFGLLLTOKGNATHDNICSGNSESTQKCGIDVT 179
Db 368 -psipatqarvcktpelpddlsdstslgtstvtctsiqegqikannskelsmcfteft 426

QY 180 -----VDKHTCPCPCAPPELLGSPVFLPPPKPDKTLMISRPEVTVCVVVDVSHE 229
Db 427 tnispgsrdpvdktchtcpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdvshe 486

QY 230 DPEVKFNWYDGVGEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP 289
Db 487 dpevkfnwydgvgevhnaaktpreeqynstyrvvsvltvlhqdwlngkeykckvsnkalp 546

QY 290 APIEKTISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGOPEN 349
Db 547 apiektiskagqpreqvyltlppsreemtknqvsllclvkgfypsdiavewesngqpen 606

QY 350 NYKTPPVLDSGDSFFLYSKLTVDKSRWQGNVFCSSVMHEALHNNHYTKSLSLSPG 406
Db 607 nykttppvlsgdsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispg 663

RESULT 11
Y80124
ID Y80124 standard; Protein; 664 AA.
XX AC Y80124;
XX DT 19-MAY-2000 (first entry)
XX DE Rat NTNR alpha and IgG fusion protein SEQ ID NO:18.
XX KW Human; neutrin receptor alpha; NTNR alpha; splenic haematopoiesis;
XX KW anaemia; thrombocytopaenia; hypoplasia; haemorrhage.
XX OS Rattus sp.
XX OS Synthetic.
XX PN US6025157-A.
XX PD 15-FEB-2000.
XX PF 24-OCT-1997; 97US-0957063.
XX PR 18-FEB-1997; 97US-0038839.
XX PR 09-JUN-1997; 97US-0049818.
XX PA (GETH) GENENTECH INC.
XX PI Hynes MA, Rosenthal A, Klein RD;
XX DR WPI; 2000-181808/16.
XX DR N-PSDB; 291460.
XX PT Isolated nucleic acid molecule encodes a neutrin receptor-alpha amino
XX PT acid sequence excluding the N-terminal signal peptide -
XX PS Claim 3; Column 97-102; 78pp; English.
XX CC The present invention describes a neutrin receptor alpha (NTNR alpha).
XX CC NTNR alpha binds neutrin. The NTNR alpha nucleic acid molecule is
XX CC useful for the expression of NTNR alpha, which may be used to identify
XX CC agonist and antagonist compounds having therapeutic applications, such

CC as enhancing splenic haematopoiesis, treating anaemia,
CC thrombocytopaenia, hypoplasia, or haemorrhage. The present sequence
CC is a NTNR alpha and IgG (immunoglobulin G) fusion protein from the
CC present invention.
XX Sequence 664 AA;

Query Match 55.2%; Score 1254.5; DB 21; Length 664;
Best Local Similarity 62.8%; Pred. No. 2e-70;
Matches 262; Conservative 30; Mismatches 56; Indels 69; Gaps 13;

QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDECLYSPVCKELQYV 72
Db 274 ncrasyrtits-cpadnyqacslgyagmifdmtpnvydsnptg---ivspwc----- 323

QY 73 KQECNRTNHRVCECKEGRYLEIEF---CLKHRSPPGFGVQAGTPERNTVCKRCPCDGF 128
Db 324 --ncrgsgnmeeeee--kflr-dftenpcl--rnaiaqfg-----ngtdvnmsspgk- 367

QY 129 FSNETSskapcrkht-----NCSVFGLLLTOKGNATHDNICSGNSESTQKCGIDVT 179
Db 368 -psipatqarvcktpelpddlsdstslgtstvtctsiqegqikannskelsmcfteft 426

QY 180 -----VDKHTCPCPCAPPELLGSPVFLPPPKPDKTLMISRPEVTVCVVVDVSHE 229
Db 427 tnispgsrdpvdktchtcpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdvshe 486

QY 230 DPEVKFNWYDGVGEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP 289
Db 487 dpevkfnwydgvgevhnaaktpreeqynstyrvvsvltvlhqdwlngkeykckvsnkalp 546

QY 290 APIEKTISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGOPEN 349
Db 547 apiektiskagqpreqvyltlppsreemtknqvsllclvkgfypsdiavewesngqpen 606

QY 350 NYKTPPVLDSGDSFFLYSKLTVDKSRWQGNVFCSSVMHEALHNNHYTKSLSLSPG 406
Db 607 nykttppvlsgdsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispg 663

RESULT 12
W70540
ID W70540 standard; Protein; 963 AA.
XX AC W70540;
XX DT 26-JAN-1999 (first entry)
XX DE Integrin beta-1 chain.
XX KW Integrin; beta-1 chain; immunoglobulin; chimeric; heterodimer complex;
XX KW inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent;
XX KW human.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..20
XX FT /note= "signal peptide"
XX FT Protein 21..963
XX FT /note= "mature protein"
XX PN W09832771-A1..
XX PD 30-JUL-1998.
XX PF 29-JAN-1998; 98WO-JP00370.
XX PF 29-AUG-1997; 97JP-0234544.
XX PR 29-JAN-1997; 97JP-0015118.
XX PA (TORA) TORAY IND INC.

XX Kainoh M, Tanaka T;
 XX WPI: 1998-427881/36.
 DR N-PSDB; V33773.
 XX Integrin-immunoglobulin chimeric protein heterodimer complexes as
 PT platelet substitutes - contain the alpha and beta integrin chains
 PT associated in stable state and bind to extracellular matrix in the
 PT presence of plasma components
 XX
 PS Claim 9; Pages 50-57; 87pp; Japanese.
 XX This represents an integrin beta-1 chain. The invention provides
 CC integrin-immunoglobulin chimeric protein heterodimer complexes that
 CC comprise an integrin alpha or beta chain associated with an
 CC immunoglobulin light or heavy chain. These chimeric proteins form
 CC heterodimer complexes, in particular with a chimeric protein containing
 CC an integrin alpha chain and an immunoglobulin chain with a chimeric
 CC protein containing an integrin beta chain and an immunoglobulin chain;
 CC the immunoglobulin chain in each case may be a heavy chain, or one of the
 CC two may be a light chain. The integrin alpha chain is preferably alpha 4
 CC or alpha 2 and the integrin beta chain is preferably beta 1. Animal cells
 CC transformed with vectors containing the DNA coding for the above chimeric
 CC proteins can be used in the preparation of the chimeric proteins and
 CC their heterodimer complexes. The heterodimer complexes, which are useful
 CC for testing potential promoters and inhibitors of the binding of
 CC integrins to their ligands, function as blood platelet substitutes and
 CC hemostatics and as diagnostic agents.
 XX
 SQ Sequence 963 AA;

Query Match 55.1%; Score 1253; DB 19; Length 963;
 Best Local Similarity 51.6%; Pred. No. 3.9e-70;
 Matches 273; Conservative 32; Mismatches 94; Indels 130; Gaps 15;
 QY 1 ETFFPKYLVDEETS--HOLLCD-----KCPPGT-----YLKQHC 33
 DB 441 dskirplgtfteeveilyceceqgsepshkchegntfecgaacrnegrgrhnc 500
 QY 34 TAKRTVACPCPDHY-----YTDWHTSDCLYCSVPKELQVVKQ-----BC 76
 DB 501 ecstdevnsedmayckrnsseicnqgcvcgckrdntneisgkfkcecdnfc 560
 QY 77 NRTN-----RVCEKRGYLETEFC--LKRSCPP-----GFGVQAGTPER 117
 DB 561 drsnglicggngvckervcecpnpytgscdcsldtstceasngqicngrgicecg 616
 QY 118 NTVKRCPCDGFSSNETSSKAP-----CRKH-----TNCVSFGLLT 153
 DB 617 --vck-ctdpxkfgqtcemcqtclgycvcaehkcvqrafnkgekktctqcsyfnitkv 673
 QY 154 QKGNATHDNI-----CSGNSE-----STQKCGI-----D 177
 DB 674 esrdklpqpvpqdpshckekdvdcwfytysvngnvmvhnvenpecptgpedpeep 733
 QY 178 VTVDKTHTCPCDPAPELLCGPSVFLPPPKPDTLMISRPETVCVVVDVSHEDPEVKFNW 237
 DB 734 kscdkthtcpcpapellgppsvflfpkpkdtlmisrtpetvcvvdvshedpevkfnw 793
 QY 238 YVDGVVHNAKTPREQYNSTRVYVTLVLDHDLNGKEYCKVSNKALPAPIEKTIS 297
 DB 794 yvdgvvhnaktprceqynstrvysvltvlhqdwlngkeyckvsnkalpapielktis 853
 QY 298 KAKGQPREQVYLPSPRDELTKNQVSLTCLVKGFPSDTAVESWNGOPENNYKTTTPV 357
 DB 854 kagqpreqvylpsrdeitknqysltclvkgfpsdiavewesngopennnykttppv 913
 QY 358 LDSGGSFFLYSKUTVDKSRQQGNVPSFSCVMHREALNHNHYTKSLSLSPG 406
 DB 914 ldsdgsfflyskitvdksrqgnvpsfscvsmhrealnhytqkslsisp 962

RESULT 13
 W73513
 ID W73513 standard; Protein; 388 AA.
 XX
 AC W73513;
 XX
 DT 02-MAR-1999 (first entry)
 XX
 DE Rabbit TGFbetaRII:Fc protein.
 XX
 KW Transforming growth factor-beta receptor; TGF-beta receptor; arthritis;
 KW fusion protein; fibroproliferative disorder; diabetic nephropathy;
 KW glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;
 KW collagen vascular disorder; therapy; rabbit.
 XX
 OS Oryctolagus sp.
 XX
 PN W09848024-A1.
 XX
 PD 29-OCT-1998.
 XX
 PF 16-APR-1998; 98WO-US07587.
 XX
 PR 18-APR-1997; 97US-0044641.
 XX
 PA (BIOI) BIOGEN INC.
 XX
 PI Cate R, Gotwals P, Koteliarsky V, Sanicola-Nadel M;
 XX
 DR WPI: 1998-609994/51.
 DR N-PSDB; V08998.
 XX
 PT Transforming growth factor-beta receptor fusion protein - used to
 PT treat fibroproliferative disorders
 XX
 PS Claim 4; Page 18-19; 70pp; English.

This sequence is a fusion protein of the rabbit transforming growth factor-beta receptor II (TGF-betaRII) and an antibody Fc region. The encoded protein is an example of a protein of the invention, which are isolated TGF-beta receptor fusion proteins that competitively inhibit binding of TGF-beta to TGF-beta receptor. The fusion protein can be used in a method for lowering the levels of TGF-beta in an individual having arthritis. It can also be used to treat medical conditions such as fibroproliferative disorders. The fibroproliferative disorder is a kidney, intraocular or pulmonary fibrosis, especially selected from diabetic nephropathy, glomerulonephritis, proliferative vitreoretinopathy, or myelofibrosis. The fusion proteins can also be used to treat collagen vascular disorders, including systemic sclerosis, polymyositis, scleroderma, dermatomyositis, and systemic lupus erythematosus. They can also be administered following coronary angioplasty, to prevent restenosis or scarring and reclosing of arteries.

Query Match 55.0%; Score 1250.5; DB 19; Length 388;
 Best Local Similarity 65.2%; Pred. No. 2e-70;
 Matches 257; Conservative 18; Mismatches 72; Indels 47; Gaps 10;
 QY 37 WKTVCAPCPDHY-----TDSWHTSDE-----CLYC---SPVCKELQVVKQECNRTN 81
 DB 17 wtiaistipphvqskvnmnmvtdngavkfqlckfcdvrsstcdnqscmnsctit-- 74
 QY 82 RVCEKRGYLETEFC-LKRSCPPGFGVQAGTPERTVCKRCPPD-----GFFSNETSS 135
 DB 75 sice-----kahevcv---avwrkndenitietvchdtklayhgfledsas 118
 QY 136 KAPCRKHTNCSVEFG---LLLTQKGNATHDNCISGNSESTQKCIDVTVDKTHTCPCPCAP 192
 DB 119 pkcimkek--kvfgctfmcscstcdndhifseeyttsspd1---vdkthtcpcpcap 173

QY	193	ELLGSPSFLPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR	252
Db	174	ellggsvfifppkpkdtlmisrtpevtcvvvdvshedpevkfnwyvdgvevhnaktkpr	233
QY	253	EEQYNSTYRVSVLTVLHGDWLNGLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLR	312
Db	234	eedynstyrvsvltvthgdwlngkeyckvsnkaipiektiskakgqprepqvylclp	293
QY	313	PSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTV	372
Db	294	psrdeltknqvsltclvkgyfypsdiavewesngqpennnyktpvpldsgsflyskltv	353
QY	373	DKSRWQGNVFCSCVMHEALHNHYTKQKLSLSLSPG	406
Db	354	dksrwqgnvfscsvmhealhnhytqkslsisp	387
RESULT 14			
Y54063	ID	Y54063 standard; Protein; 388 AA.	
XX	XX	Y54063;	
XX	XX		
XX	DT	27-MAR-2000 (first entry)	
XX	XX	Amino acid sequence of TGF-beta type II receptor variant/IgG1 fusion.	
DE	DE		
KW	KW	Splice variant; rabbit; transforming growth factor-beta; TGF-beta;	
KW	KW	type II receptor; Fc portion; human IgG1; fusion protein; arthritis;	
KW	KW	fibroproliferative disease; renal; intra-ocular; pulmonary; fibrosis;	
KW	KW	diabetic nephropathy; glomerulonephritis; collagen vascular disease;	
KW	KW	proliferative vitreoretinopathy; myelofibrosis; systemic sclerosis;	
KW	KW	polymyositis; scleroderma; dermatomyositis; systemic lupus erythematosus;	
KW	KW	tenositis; wound; connective tissue production; adhesion; scarring;	
KW	KW	post-radiation fibrosis.	
XX	XX		
OS	OS	Synthetic.	
OS	OS	Oryctolagus cuniculus.	
OS	OS	Homo sapiens.	
XX	XX	WC09965948-Al.	
XX	XX		
XX	PD	23-DEC-1999.	
XX	XX		
PF	PF	16-JUN-1999; 99WO-US13629.	
XX	XX		
PR	PR	16-JUN-1998; 98US-0089452.	
XX	XX		
PA	PA	(BIOJ) BIOGEN INC.	
XX	XX		
PI	PI	Koteliansky V, Gotwals P, Cate R, Sanicola-Nadel M;	
XX	XX		
DR	DR	WPI; 2000-106083/09.	
DR	DR	N-PSDB; Z45251, Z45252.	
XX	XX		
PT	PT	New fusion protein of a splice variant of transforming growth	
PT	PT	factor-beta receptor, for inhibiting the growth factor, e.g. in	
XX	XX	treatment of fibrosis	
XX	XX		
PS	PS	Disclosure; Page 62-63; 69pp; English.	
XX	XX		
CC	CC	The present sequence represents a splice variant of a rabbit transforming	
CC	CC	growth factor-beta (TGF-beta) type II receptor fused to the Fc portion	
CC	CC	of human IgG1. The fusion protein has higher affinity for TGF-beta than	
CC	CC	fusion proteins comprising the non-variant form of the receptor. The	
CC	CC	fusion protein contains soluble TGF-beta receptor constructs that are	
CC	CC	devoid of a transmembrane region (and are secreted from the cell) but	
CC	CC	retain the ability to bind TGF-beta. The protein competitively inhibits	
CC	CC	binding of TGF-beta to cellular receptors and/or forms an inactive	
CC	CC	complex with TGF-beta. The protein is used to reduce levels of TGF-beta,	
CC	CC	for treatment of arthritic conditions associated with overexpression	
CC	CC	of TGF-beta, especially fibroproliferative diseases, e.g. renal,	
CC	CC	intra-ocular or pulmonary fibrosis; diabetic nephropathy;	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:20:09 ; Search time 135.68 Seconds
(without alignments)
348.998 Million cell updates/sec

Title: US-09-389-782A-7
Perfect score: 2264
Sequence: 1 ETEPPKYLHYDEETSHQLL.....VMHEALHNHYTKSLSISPG 404

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_15:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1013	44.7	372	4 Q9UHP4	Q9uHP4 homo sapien
2	1013	44.7	401	4 O00300	O00300 homo sapien
3	892	39.4	401	11 O08712	O08712 mus musculus
4	879	38.8	401	11 O08727	O08727 rattus norv
5	819.5	36.2	437	11 Q9RIA4	Q9ria4 mus musculus
6	433.5	19.1	300	4 Q95407	Q95407 homo sapien
7	406.5	18.0	302	13 Q9PUS0	Q9pus0 salvelinus
8	337.5	14.9	439	4 Q16042	Q16042 homo sapien
9	335	14.8	459	11 Q62327	Q62327 mus musculus
10	328	14.5	482	11 O88734	O88734 mus musculus
11	298	13.2	655	4 Q75509	Q75509 homo sapien
12	269.5	11.9	416	4 Q9NPP6	Q9npp6 homo sapien
13	269.5	11.9	684	13 Q90544	Q90544 ginglymosto
14	268.5	11.9	384	4 Q9UP60	Q9up60 homo sapien
15	253.5	11.2	625	11 Q35305	Q35305 mus musculus
16	252.5	11.2	616	4 Q9Y6Q6	Q9y6q6 homo sapien
17	239	10.6	349	12 O57099	O57099 monkeypox v
18	237	10.5	349	12 O57291	O57291 monkeypox v
19	237	10.5	349	12 O57100	O57100 monkeypox v

20	237	10.5	349	12	O57101	monkeypox v
21	237	10.5	349	12	O57102	monkeypox v
22	233.5	10.3	348	12	O57277	monkeypox v
23	233.5	10.3	348	12	O57103	monkeypox v
24	233.5	10.3	348	12	O57108	monkeypox v
25	229	10.1	348	12	O57112	variola vir
26	229	10.1	348	12	O85407	variola vir
27	226.5	10.0	349	12	O57284	camelopox vi
28	226.5	10.0	349	12	O57098	camelopox vi
29	226	10.0	349	12	O57110	variola vir
30	226	10.0	349	12	O57111	variola vir
31	226	10.0	349	12	O89118	variola vir
32	226	10.0	349	12	O89098	variola vir
33	223	9.8	350	12	O57116	cowpox viru
34	222.5	9.8	349	12	O57097	camelopox vi
35	222.5	9.8	355	12	O85308	cowpox viru
36	221	9.8	349	12	O57109	variola vir
37	220.5	9.7	349	12	O57305	cowpox viru
38	217.5	9.6	326	12	O57120	cowpox viru
39	217.5	9.6	326	12	O57122	cowpox viru
40	217.5	9.6	351	12	O73559	cowpox viru
41	217.5	9.6	360	12	O57118	cowpox viru
42	216.5	9.6	351	12	O57117	cowpox viru
43	214.5	9.5	350	12	O57123	cowpox viru
44	212.5	9.4	347	12	O57115	cowpox viru
45	209.5	9.3	283	4	Q92956	homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	372 AA.
Q9UHP4	AC Q9UHP4;			
ID	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	OSTEOPROTEGERIN (FRAGMENT).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PLACENTA;			
RA	He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;			
RT	"Cloning and Expression of Osteoprotegerin from Homo sapiens.";			
RL	Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).			
DR	EMBL; AF134187; AAF20168.1;			
DR	HSSP; P25942; 1CDF.			
DR	INTERPRO; IPR001368;			
DR	PFAM; PF00020; TNFR_c6; 3.			
DR	PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.			
DR	PROSITE; PS50050; TNFR_NGFR_2; 2.			
FT	NON_TER 1			
SQ	SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;			

Query Match 44.7%; Score 1013; DB 4; Length 372;
Best Local Similarity 99.4%; Pred. No. 2.4e-81;
Matches 174; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ETFPKYLHYDEETSHQLLCKPCPGTYLKQHCATKWTVCAPCPDHYTDSWHTSDECL	60
Db	1	ETFPKYLHYDEETSHQLLCKPCPGTYLKQHCATKWTVCAPCPDHYTDSWHTSDECL	60
QY	61	YCSFVKELQYVQECNRTNHRVCECKEGRYLIEFCLKHRSCPPGFGVYVQAGTPERTV	120
Db	61	YCSFVKELQYVQECNRTNHRVCECKEGRYLIEFCLKHRSCPPGFGVYVQAGTPERTV	120
QY	121	CKRCPDGFSSNETSSKAPCRKHTNCVSFGLLLTQKGNATHDNCISGNSSTQKSG	175

2

RESULT

000300 PRELIMINARY; PRT; 401 AA.

ID 000300 PRELIMINARY; PRT; 401 AA.

AC 000300; 060236;

AT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11B).

GN TNFRSF11B OR OPG OR OCIF.

DE Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

[1]

SEQUENCE FROM N.A.

PC TISSUE=KIDNEY;

RC MEDLINE=97262071; PubMed=9108485;

RX Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,

RA Luethy R., Nguyen H.O., Wooden S., Bennett L., Boone T., Shimamoto G.,

RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,

RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,

RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,

RA Suggs S., Boyle W.J.;

RA "Osteoprotegerin: a novel secreted protein involved in the regulation

RT of bone density.;"

RL Cell 89:309-319(1997).

RL [1]

SEQUENCE FROM N.A.

PC TISSUE=LUNG FIBROBLAST;

RC MEDLINE=98151033; PubMed=9492069;

RX Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,

RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,

RA Tsuda E., Morinaga T., Higashio K.;

RA "Identity of osteoclastogenesis inhibitory factor (OCIF) and

RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits

RT osteoclastogenesis in vitro.;"

RT Endocrinology 139:1329-1337(1998).

RN [3]

SEQUENCE FROM N.A.

PC TISSUE=PLACENTA;

RC MEDLINE=98351569; PubMed=9688283;

RX Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;

RA "Cloning and characterization of the gene encoding human

RT osteoprotegerin/osteoclastogenesis-inhibitory factor.;"

RT Eur. J. Biochem. 254:685-691(1998).

CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES

CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY

CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO

CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN

CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.

CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY

CC SIMILARITY).

CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART, IN

CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN

CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN

CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.

CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

DR EMBL; AB0021146; BAA25910.1;

DR EMBL; AB008822; BAA32076.1;

DR EMBL; AB008821; BAA32076.1; JOINED.

DR EMBL; U94332; AAB53709.1;

DR HSSP; P25942; 1CDF.

DR MIM; 602643;

DR INTERPRO; IPR001368;

DR PFAM; PF00020; TNFR_C6; 3.

DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

DR PROSITE; PS50050; TNFR_NGFR_2; 2.

DR PRODOM; PD000771; -; 1.

RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,
 RA Higashio K.;
 RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
 RL gene and its expression in embryogenesis.";
 CC Gene 215:339-343(1998).
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
 CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
 CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
 CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
 CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,
 CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND
 CC PLACENTA. NOT DETECTED IN SPLEEN.
 CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT
 CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY
 CC 15 TO DAY 17.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC EMBL; U94331; AAB53708.1; -.
 CC EMBL; AB013898; BAA28269.1; -.
 CC EMBL; AB013903; BAA33388.1; -.
 CC EMBL; AB013899; BAA33388.1; JOINED.
 CC EMBL; AB013900; BAA33388.1; JOINED.
 CC EMBL; AB013901; BAA33388.1; JOINED.
 CC EMBL; AB013902; BAA33388.1; JOINED.
 CC HSP; P25942; 1CDF.
 CC MGD; MG1:109587; Opg.
 CC INTERPRO; IPR000488; -.
 CC INTERPRO; IPR001368; -.
 CC PFAM; PF00020; TNFR_C6; 3.
 CC PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 CC PROSITE; PS0017; DEATH_DOMAIN; 1.
 CC PROSITE; PS50050; TNFR_NGFR_2; 2.
 CC PRODOM; PD000771; -. 1.
 KW Glycoprotein; Repeat; Cytokine; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 401 OSTEOPROTEGERIN.
 FT DOMAIN 23 201 4 X TNFR-CYS.
 FT REPEAT 23 63 TNFR-CYS 1.
 FT REPEAT 64 106 TNFR-CYS 2.
 FT REPEAT 107 143 TNFR-CYS 3.
 FT REPEAT 144 201 TNFR-CYS 4.
 FT DOMAIN 306 365 DEATH DOMAIN.
 FT DISULFID 41 54 BY SIMILARITY.
 FT DISULFID 44 62 BY SIMILARITY.
 FT DISULFID 65 80 BY SIMILARITY.
 FT DISULFID 83 97 BY SIMILARITY.
 FT DISULFID 118 142 BY SIMILARITY.
 FT DISULFID 145 160 BY SIMILARITY.
 FT CARBOHYD 165 178 POTENTIAL.
 FT CARBOHYD 178 178 POTENTIAL.
 FT CARBOHYD 289 289 POTENTIAL.
 FT VARIANT 138 138
 FT VARIANT 138 138
 FT VARIANT 161 161
 FT VARIANT 161 161 I -> R (IN STRAINS 129/OLA AND NIH
 FT VARIANT 165 165 SWISS).
 FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH
 FT VARIANT 288 288 SWISS).
 FT VARIANT 288 288 S -> A (IN STRAINS 129/OLA AND NIH
 FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH
 FT VARIANT 296 296 SWISS).
 SQ SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;

Query Match 39.4%; Score 892; DB 11; Length 401;
 Best Local Similarity 86.9%; Pred. No. 1.2e-70;
 Matches 152; Conservative 7; Mismatches 16; Indels 0; Gaps 0;
 QY 1 ETPPKYLVDETSHTLLCDKCPPTYLKQHCTAKWKTVCAPCPDHYTDSWHTSDECL 60

Db 22 ETLPPKLVHDPETGHTLLCDKCPPTYLKQHCTVRRKTLVPCPDHVSYDSWHTSDECL 81
 QY 61 YCSPVKELQYQKCNRTHNVCECKEGRYLEIEFCLKHSRCPGPGVGVQAGTPERNIV 120
 Db 82 YCSPVKELQYQKCNRTHNVCECKEGRYLEIEFCLKHSRCPGPGVGVQAGTPERNIV 141
 QY 121 CKRCPDGFESNTSKAPCRKHTNCSVFLGTLTKGNATHDNICSNSESTQKSG 175
 Db 142 CKRCPDGFESNTSKAPCRKHTNCSVFLGTLTKGNATHDNVCSGNREATQKCG 196
 RESULT 4
 O08727 PRELIMINARY; PRT; 401 AA.
 ID O08727
 AC O08727
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
 DE (OCIF).
 GN TNFRSF11B OR OPG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
 OC NCBI_TaxID-10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=INTESTINE;
 RX MEDLINE-97262071; PubMed-9108485;
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
 RA Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,
 RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
 RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
 RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
 RA Suggs S., Boyle W.J.;
 RT "Osteoprotegerin: a novel secreted protein involved in the regulation
 of bone density";
 RL Cell 89:309-319(1997).
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
 CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
 CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
 CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
 CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC EMBL; U94330; AAB53707.1; -.
 CC HSP; P25942; 1CDF.
 CC INTERPRO; IPR001368; -.
 CC PFAM; PF00020; TNFR_C6; 4.
 CC PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 CC PROSITE; PS50050; TNFR_NGFR_2; 2.
 CC PRODOM; PD000771; -. 1.
 KW Glycoprotein; Repeat; Cytokine; Signal.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT CHAIN 22 401 OSTEOPROTEGERIN.
 FT DOMAIN 23 201 4 X TNFR-CYS.
 FT REPEAT 23 63 TNFR-CYS 1.
 FT REPEAT 64 106 TNFR-CYS 2.
 FT REPEAT 107 143 TNFR-CYS 3.
 FT REPEAT 144 201 TNFR-CYS 4.
 FT DOMAIN 306 365 DEATH DOMAIN.
 FT DISULFID 41 54 BY SIMILARITY.
 FT DISULFID 44 62 BY SIMILARITY.
 FT DISULFID 65 80 BY SIMILARITY.
 FT DISULFID 83 97 BY SIMILARITY.
 FT DISULFID 87 105 BY SIMILARITY.
 FT DISULFID 118 142 BY SIMILARITY.
 FT DISULFID 145 160 BY SIMILARITY.
 FT CARBOHYD 98 98 POTENTIAL.
 FT CARBOHYD 165 165 POTENTIAL.

FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
SQ SEQUENCE 401 AA; 46192 MW; FEC6A31F1D4E573A CRC64;

Query Match 38.8%; Score 879; DB 11; Length 401;
Best Local Similarity 85.1%; Pred. No. 1.7e-69;
Matches 149; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 ETPEPKYLHYDETHSHOLLCDKCPGTYLKHCHTAKWKTVCAPCPDHYHDTSDHWTSDCL 60
DB 22 ETPEPKYLHYDETHSHOLLCDKCPGTYLKHCHTAKWKTVCAPCPDHYHDTSDHWTSDCL 81
QY 61 YCSPVKELQVQKQECNTHNRVCECKEGRYLEFELCLKHRSCTPGGTVQAGTPERTV 120
DB 82 YCSPVKELQVQKQECNTHNRVCECKEGRYLEFELCLKHRSCTPGGTVQAGTPERTV 141
QY 121 CKRCPDGFNETSSKAPCRKHTNCVFGLLLTQKGNATHDNCISGSESTQKSG 175
DB 142 CKRCPDGFNETSSKAPCRKHTNCVFGLLLTQKGNATHDNCISGSESTQKSG 196

RESULT 5

Q9RIA4 PRELIMINARY; PRT; 437 AA.

AC Q9RIA4; MEDLINE-99087326; PubMed-9872321;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SQ SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv)."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -;
DR HSP; P01842; 7FAB.
DR INTERPRO; IPR003006; -;
DR PFAM; PF00047; ig; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 36.2%; Score 819.5; DB 11; Length 437;
Best Local Similarity 62.8%; Pred. No. 3.2e-64;
Matches 140; Conservative 43; Mismatches 35; Indels 5; Gaps 2;

QY 184 CPP--CPAPELLGSPSVLEFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYDQVE 241
DB 217 KCPCICITPEV---SSVFIFPPKPKDVLITLTPKVCVVVDISKDDPEVQFSFWDVE 273
QY 242 VHNAKTPREQYNSTYRVYSLTVLHODWLNGKEYCKYKSNKALPAPIETKISKAGQP 301
DB 274 VHTAQTPREQYNSTYRVYSLTVLHODWLNGKEYCKYKSNKALPAPIETKISKAGQP 333
QY 302 REPOVYTLPSRDELTKNOYSLCLVGFYPSDITAVESNGQPENNYKTPPVLDSDGS 361
DB 334 KAPOVYTIPTPKQEKQADKSLVLCMTITDFPEDITVEMQNGQPAENYKNTQPIMDTDS 393
QY 362 FFYLSKLTVDKSRWQGNVFSVCSVMHEALHNHYTKSLSPG 404
DB 394 YFVYSLNVQKSNWEACNTTCSVLHESLHNHTEKLSHSPG 436

RESULT 6

Q95407 PRELIMINARY; PRT; 300 AA.
AC Q95407; MEDLINE-99087326; PubMed-9872321;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN DCR3 OR TR6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SQ SEQUENCE FROM N.A.
RA Pitti R.M., Masters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA Goddard A.D., Botstein D., Ashkenazi A.;
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
RT colon cancer."
RL Nature 396:699-703(1998).
RN [2]
SQ SEQUENCE FROM N.A.
RA Tissue-Blood;
RC Tissue-Blood;
RX Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT "A newly identified member of tumor necrosis factor receptor
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis."
RL J. Biol. Chem. 274:13733-13736(1999).
RN [3]
SQ SEQUENCE FROM N.A.
RC Tissue-Pancreas;
RX MEDLINE-20122600; PubMed-10655513;
RA Bai C., Connolly B., Metzger M.L., Hillard C.A., Liu X., Sandig V.,
RA Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RT cluster."
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR EMBL; AF104419; AAD03056.1; -;
DR EMBL; AF134240; AAD29688.1; -;
DR EMBL; AF217796; AAF35244.1; -;
DR EMBL; AF217793; AAF33685.1; -;
DR EMBL; AF217794; AAF33686.1; -;
DR HSP; P25942; 1CDF.
DR INTERPRO; IPR000561; -;
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AE33718449AF CRC64;

Query Match 19.1%; Score 433.5; DB 4; Length 300;
Best Local Similarity 42.0%; Pred. No. 2.1e-30;
Matches 73; Conservative 30; Mismatches 68; Indels 3; Gaps 1;

QY 5 PKYLHYDETHSHOLLCDKCPGTYLKHCHTAKWKTVCAPCPDHYHDTSDHWTSDCLYCSP 64
DB 34 PTPWRDAETGERLVCAQCPCPGTFVQPCRDSPPTTCGCPPHRYTGFNWLRCRCNV 93
QY 65 VKCELQVYKQECNTHNRVCECKEGRYLEFELCLKHRSCTPGGTVQAGTPERTVCKRC 124
DB 94 LCGEREEARACHATHNRACRCRTGFFAHAGFCLEHASCPCPGAGVIAPGTPSQNTQCP 153
QY 125 PDGFFSNETSSKAPCRKHTNCVFGLLLTQKGNATHDNI---CSGNSESTQKSG 175
DB 154 PPGTFSSSSSSSQCPQPHNRCTALGLALNVPGSSSHDTLCTCTGFPPLRVPG 207

```

RESULT 7
Q9PUSO PRELIMINARY; PRT; 302 AA.
ID Q9PUSO
AC Q9PUSO
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE DCOY RECEPTOR.
OS Salvelinus fontinalis (Brook trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
RA Bobe J., Goetz F.W.;
RT "A tumor necrosis factor receptor homolog is up-regulated in the brook
trout (Salvelinus fontinalis) ovary at the completion of ovulation.";
RL Biol. Reprod. 0:0-0(1999).
DR EMBL; AF156738; AAD56428.1; -.
DR HSSP; P19438; 1EXT.
DR INTERPRO; IPR000561; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 18.0%; Score 406.5; DB 13; Length 302;
Best Local Similarity 43.1%; Pred. No. 5e-28;
Matches 72; Conservative 32; Mismatches 60; Indels 3; Gaps 2;

QY 11 DEETSHOLLCDKCPGGTYLKQHCCTAKWTVACPDHYTDSWHTSDECLYSPVCKELQ 70
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
27 DSGSLIVDCRPPGGTYLRAPCSAMKSDCAECPNGAYTEFNHISKLRCS-MCAENQ 85

QY 71 YVQECNTHNRCVCEKGRYLE--EIFELKHRSCPPGFGVQAGTPERNVCKRCPDGF 128
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
86 VVQECSPSNCECEKEGYFYNKKYACIKHKCECPGPGYANTTGTGPHQDTECVQCOAGF 145

QY 129 FSNETSKAPCRKHTNCSVFGLLLTQGNATHDNICSGNSESTOKSG 175
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
146 YSEVSSAKATCLAQSNCKVGLRVVLKQDWHNTLCASCYDLKTRDG 192

RESULT 8
Q16042 PRELIMINARY; PRT; 439 AA.
ID Q16042
AC Q16042
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91370650; PubMed=1966549;
RA Dembic Z., Loeschner H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,
RA Brockhaus M., Lesslauer W.;
RT "Two human TNF receptors have similar extracellular, but distinct
intracellular, domain sequences.";
RL Cytokine 2:231-237(1990).
DR EMBL; S63368; AAB19824.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.

```

```

DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
DR PRODOM; PD000771; -.
SQ SEQUENCE 439 AA; 46090 MW; FEBCEB329CC67FE6 CRC64;

Query Match 14.9%; Score 337.5; DB 4; Length 439;
Best Local Similarity 26.6%; Pred. No. 9.4e-22;
Matches 107; Conservative 44; Mismatches 163; Indels 89; Gaps 15;

QY 9 HYDEETSHOLLCDKCPGGTYLKQHCCTAKWTVACPDHYTDSWHTSDECLYSPVCKE 68
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
23 YDQ--TAQCCSKCSPGQHAQVCTKTSDTVCDCSDSTYTQLMNVWPECLSCGSRCS 80

QY 69 LOYVVKQECNTHNRCVCEKGRYLEI-----EFLKHRSCPPGFGVQAGTPERNVCK 122
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
81 DQVETQACTREQNRICTCRPGWYCALSKQECRLCAPLKRCPGFGVARPGTETSDVVVK 140

QY 123 RCPDGFESNETSSKAPCRKHTNCSVFGLLLTQGNATHDNICSGNSESTOKSGG----- 177
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
141 PCAPGTFSTNTSSDTCRPHQICNVAI-----PCNASMDAVCTSTSPTRSNAPGAVHLPQ 196

QY 178 -GGGCTCPPCPAPE-----LLGSPSVFLPPPKPKDT-----LMISRT----- 215
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
197 PVSTRSQHTOPTPEPSTAPSTFLLPMGPS-----PPAEGSTGDFALPVGLIVGTALGL 252

QY 216 ---EVTCCVVVDVSHEDP-----EVKFNWYVDGVEVHNK-TKPREQYNSTYRVVSVLT 266
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
253 IGVVNCVIMTVQKKKPLCLQREAK----VPHLPADKARGTGGPEQOH-----LLIT 300

QY 267 LHODWLNKGEYKCKVSNKALPAPIETKISKAKGPREPOVYVTPPSRDELTK----- 318
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 APSSSSSLSASSALDRAP-----TRNQPPQPGVEASGAGERASTGSSDSPG 351

QY 319 ---NQVSLTCLVKGFPDIAVEWESN-----GQPENNYKTP 353
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 GHGTQVNVTCIVNVCSSSDHSSQCSQASSTMGTDTSPPSP 394

RESULT 9
Q62327 PRELIMINARY; PRT; 459 AA.
ID Q62327
AC Q62327
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=95178848; PubMed=7873884;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
gene.";
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.
DR HSSP; P19438; 1NCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
FT NON_TER 1
FT VARIANT 87 93 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

```


Db 54 YRVDNRATGVLCDKCPAGTYVSEHCTNLSRVCSGPCVGTETRHENGIEKCHDCSQPC 113
QY 67 KELQYVQECNRTHNRVCECKEGRYLEIEFCLAKHRSCPPGFGVVOAGTPEPNTVCKRCPD 126
Db 114 PWFMIKLPALTDRECTCPGFMFOSNATCAHTVCPVGWGRKKGTETEDVRCKOCAR 173
QY 127 GFTSNSTSSKAPCKRHTNCSVFLGILLTKGNATHDNICSGNSESTOKSGGGGGGTCTP- 185
Db 174 GTFSDVPSSVMKCAVTDCLSLQNLVLIKPGTKETDNVC-----GTLPS 216
QY 186 -----PCPAPELLGSPVLEFPKPKDTLMSRTPETVCVVVDVSHED-----PEVKF 233
Db 217 FSSSTSPSP-----GTAIF-----PRP-EHMETHEVPSSTIYVPGKMNSTESNASVRPKV-L 267
QY 234 NWVVDGVEVHNATKPREQYNSTYRVSVTLVTHODWLNGKEYCKVCKSNKALPAPI--- 290
Db 268 SSTQEGTVPNTSSARGKEDVNT--LPNLQVNVHQ---QGPHRHIL--KLLPSMEATG 320
QY 291 -EKTIISKAKQPR 302
Db 321 GEKSTPIKPKR 333

RESULT 12

Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Aulfray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA Leirach H., Poustka A., Lundeberg J.;
RT "The European IMAGE consortium for integrated Molecular analysis of
RT human gene transcripts";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389978; CAB97534.1; -
FT NON_TER
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 11.9%; Score 269.5; DB 4; Length 416;
Best Local Similarity 25.8%; Pred. No. 8.6e-16;
Matches 90; Conservative 42; Mismatches 136; Indels 81; Gaps 13;
QY 104 PPGFGVVOAGTPEPNTVCKRCPDGFPSNETSSKAPCKRHTNCSVFLGILLTKGNATHDN 162
Db 82 PKVFPILSDTPDGNVNVACLVQGFPPQEP-----LSVTWSESGQN 123
QY 163 ICSGNSESTOKSGG-----GGGGGTC-----PPCPAPELL 193
Db 124 VTARNFPPSDASGDLTYTTSSQLTLPATQCPDGKSVTCHVKHTNPSQDVTVPVCPV--- 180
QY 194 GGSVFLFPKP-----KDTLMSRTPETVCVVVDVSHEDPEVKFNWYVDGVE 241
Db 181 -----PPPCCHPLSLHRLPALEDLLGSEANLTCTLGL--RDASGATFTVPSGK 231
QY 242 VHNATKPREQYNSTYRVSVTLVTHODWLNGKEYCKVCKSNKALPAPIEKTISKAKGP 301
Db 232 --SAVGOPPRDLICGYSVSVLPGCAQPNWGHETFTCTAAHPELKTPLPANITKS--GNT 288
QY 302 REPQVYTLPPSRDELTKNQ-VSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTTPPVLD- 357

Db 289 FRPEVHLLPPSPSEALNELVTLCTLARGFSPKDVLRWLQSQDELPREKYLTLTASQEP 348
QY 358 SDG--SFFLYSKLTVDKSRVQOQNVFSCSVNHEALHNHYTKQSLSLSPG 404
Db 349 SQGTTTFAVTSILRVAADKKGDTFSCVMGHEALPLAFTQKTIDRLAG 397

RESULT 13

Q90544 PRELIMINARY; PRT; 684 AA.
AC Q90544;
DT 01-NOV-1996 (TREMBLrel. 01; Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RX MEDLINE=95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks";
RL Nature 374:168-173(1995).
DR EMBL; U18701; AAB48195.1; -
DR HSSP; P01857; 1FC1.
DR INTERPRO: IPR003006; -
DR PFAM; PF00047; ig; 6.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Signal.
FT SIGNAL
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DEF CRC64;

Query Match 11.9%; Score 269.5; DB 13; Length 684;
Best Local Similarity 26.2%; Pred. No. 1.6e-15;
Matches 100; Conservative 53; Mismatches 162; Indels 67; Gaps 16;
QY 51 DSHHTSDECLYCS-----PVCKELOYVKQECNRTHNRVCECKEGRYLEIEFCLKHS 102
Db 315 EEWQSGVE-YTCSAKQDQSTPVYKTRKARVEPTKPHRL----- 354
QY 103 CPPGFGVVOAGTPEPNTVCKRCPDGFPSNETS---SKAPCKRHTNCSVFLGILLTKGNAT 159
Db 355 LPSPPEIQSTSATLTCLLR---GFYPDKVSVSWQKDDVSVSANVTFPTALQ--DLT 409
QY 160 HDNICSGNSESTQ-KSGGGGGGGTCTP-----CPAPELLGSPVFLFP 203
Db 410 FSTRSLNLTAWEKSGAKYCTCTASHPPSPQSTVKVRIRNQKVDRCOTDI---SVSLKP 465
QY 204 KPQDTLMISTPEVTCVVVDVSHEDDP-VKFNWYVDGVEVHNATKPREQYNSTYRVVS 262
Db 466 -PFEIWTQQTATVCEIV---YSDLENIRKVFQVNGVERKKGVETQNPWMSGSKSTIVS 521
QY 263 VLVVTHODWLNGKEYCKVCKSNKALPAPIEKTISKAK-GQPREQVYTLPPSRDEL-TKNQ 320
Db 522 KLVKMASEWDSGIEYVCLVEDSELTPTPKASIRKANVYQMHPPKVVLLHSTDEIDTENS 581
QY 321 VSLTCLVKGFYPSDIAVEWESNGQ-PENNYKTTTPPVLDSDGSFFLYSKLAVDKSRWQGN 379
Db 582 ATMLCLATNFHPAEIYVGMWANDTLSDGYRTQVDSEKSGSSEFTDRLRLTAEWNSDT 641
QY 380 VFCSSVMHEALHN---HYTOKS 398
Db 642 TYSCLVGHPSLNRDLIRSTNKS 663

```
RESULT 14
O9UP60 PRELIMINARY; PRT; 384 AA.
AC O9UP60;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE SN73 PROTEIN.
GN SN73.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;
RA "Identification and characterization of SNC73, a gene which is down-
RA regulated in colorectal cancer.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067420; AAC19365.1;
DR HSSP; P01825; 7FAB
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 19; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;

Query Match 11.9%; Score 268.5; DB 4; Length 384;
Best Local Similarity 26.2%; Pred. No. 9.6e-16;
Matches 89; Conservative 40; Mismatches 138; Indels 73; Gaps 12;

Qy 115 PERNTVRCRCPDGFPSNETSSKAPCRKHTNCVFLGLLTQKGNATHDNCSESTQKS 174
Db 49 PDGNVVIACLVQGFPPQEP-----LSVTWSESGQGVGTARNPPSODA 90

Qy 175 GGG-----GGGGTGC-----PPCPAPPELLGGPSVFLFPPKP 205
Db 91 SGDLTSSQLTLPATQCLAGKSVTCHVKYHTNPQSDVTVCVPVSTPTPTSPST-PTTP 149

Qy 206 K-----DTLMISRPTEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPR 250
Db 150 SPSCCHPRLSLHRPALEDDLLGSEANLCTLTGL-RTDASGVFTFTWTPSSGK--SAVQGGP 206

Qy 251 EEQYNSTYRVYSVLVTLVHODWLNQKEYCKVSKNALPAPIEKTISKAKGQPREPOVYTL 310
Db 207 ERDLCCGYSVSVLPGCCAEPNHHGKTFTCTAAYPESKTPLTATLSKS-GNTRFPEVHLTP 265

Qy 311 PSRDELTKNQ-VSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTTPPVLD-SDG--SFPL 364
Db 266 PSEELALNELVTLTCLARGFSPKDLVRLQGSQGLPREKYLTLWASRQEPSQGGTTTFAV 325

Qy 365 YSKLTVDKSRWQGNVFCSVNWEALHHNYTKSLSLSPG 404
Db 326 TSILRVAEDMKGDTFCWVGHEALPLAFTQKTIDRLAG 365

RESULT 15
O35305 PRELIMINARY; PRT; 625 AA.
AC O35305;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE RECEPTOR ACTIVATOR OF NF-KAPPA-B PRECURSOR (TNF-RELATED ACTIVATION-
DE INDUCED CYTOKINE RECEPTOR) (RANK).
GN TNFRSF11A OR RANK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE=LIVER EPITHELIUM;
RX MEDLINE=98032977; PubMed=9367155;
RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
RA Tometsko M.E., Roux E.R., Teepe M.C., DuRose R.F., Cosman D.,
RA Galibert L.;
RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
RT and dendritic-cell function.";
RL Nature 390:175-179(1997).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS RANKL.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: CONTAINS A LA-NGR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; AF019046; AAB86810.1;
DR HSSP; P25942; 1CDF.
DR MGD; MGI:1314891; Tnfrsf11a.
DR INTERPRO; IPR000561;
DR INTERPRO; IPR001368;
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGRF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGRF_2; 1.
DR PRODOM; PD000771; 1.
KW Receptor; Glycoprotein; Transmembrane; Signal; Repeat.
FT SIGNAL 1 30
FT CHAIN 31 625
FT DOMAIN 31 214
FT TRANSMEM 215 235
FT DOMAIN 236 625
FT DOMAIN 34 196
FT REPEAT 34 70
FT REPEAT 71 114
FT REPEAT 115 153
FT REPEAT 154 196
FT DISULFID 35 47
FT DISULFID 48 61
FT DISULFID 51 69
FT DISULFID 72 87
FT DISULFID 93 113
FT DISULFID 115 125
FT DISULFID 127 134
FT DISULFID 128 152
FT DISULFID 155 170
FT DISULFID 176 195
FT CARBOHYD 106 106
FT CARBOHYD 175 175
SQ SEQUENCE 625 AA; 66621 MW; F8C1872E99511D8E CRC64;

Query Match 11.2%; Score 253.5; DB 11; Length 625;
Best Local Similarity 25.3%; Pred. No. 3.6e-14;
Matches 103; Conservative 41; Mismatches 146; Indels 117; Gaps 18;

Qy 20 CDKCPPTGLYKHCTAKWKTVCAPCPDHYVDSWHTSDECLYCSPVC---KELQVVKQEC 76
Db 48 CSRCEPGKYLSSKCTPTSDSVCLPCGPDPEYLDTWNEEDKCL-LHKVCDAGKALVAV-DPG 105

Qy 77 NRTNRVCECKEGRY--LEIEFLKHKRSCPPGFVGVQAGTPERNTVCKRCPDGFESNETS 134
Db 106 NHTAPRCACACTAGYHNSDCECCRRNTECAPGFGAQAHPQLQNKDVTCTPCLLGFSDVFS 165

Qy 135 SKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQKSGGGGGGGTCTPPCPAPELLG 194
Db 166 STDCKPWTNCTLLGKLEAHQGTTESDVVCS--SSMTLRR-----PPKEAQAYL- 212

Qy 195 GPSVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPREQY 254
Db 213 -PSITVL-----LLFISVVVVAIIIFGV-----YVRKGGK----- 241

Qy 255 NSTYRVSVLTVLHODWLN-----GKEY---KCKVSNKA-----LPADIE 291
Db 242 -----ALTANLWNVNDACSSLGKNKSSGDRCSGSHSATSSQOEVCCEGILLMTREE 293

Qy 292 KTIK-----AKGQP-----REPOVYTL-----PSRDELTKNQVTS 322
```


Db 294 KMVPEDGAGVCGPVCAAGGWAERVDRSRTFTLVSEVETQGDLSRKIPTEDXYTDRPSQPS 353
Qy 323 LTCIV-----KGFYPSDIAVWESNGQPENNYKTTTPPVLDSDGSFF 363
Db 354 TGSLLLIQOGSKSIPFPQEPLEVGENDSLQCFGTGTSTVDSEGCDP 400

Search completed: March 1, 2001, 09:20:11
Job time: 409 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:47 ; Search time 40.97 Seconds
(without alignments)
318.448 Million cell updates/sec

Title: US-09-389-782A-7
Perfect score: 2264
Sequence: 1 ETEPPKYLHYDEETSHQLC.....VMHEALHNYTQKSLSPG 404

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	ID	Description
1	1204	53.2	330	1 GC1_HUMAN	P01857 homo sapien
2	1137.5	50.2	326	1 GC2_HUMAN	P01859 homo sapien
3	1133	50.0	327	1 GC4_HUMAN	P01861 homo sapien
4	1125.5	49.7	290	1 GC3_HUMAN	P01860 homo sapien
5	909	40.2	323	1 GC_RABIT	P01870 oryctolagus
6	890.5	39.3	329	1 GC2_CAVPO	P01862 cavia porce
7	841.5	37.2	329	1 GC3_MOUSE	P22436 mus musculus
8	840	37.1	333	1 GCB_RAT	P20761 rattus norv
9	835.5	36.9	398	1 GC3M_MOUSE	P03987 mus musculus
10	824.5	36.4	326	1 GC1_RAT	P20759 rattus norv
11	812.5	35.9	324	1 GC1_MOUSE	P01868 mus musculus
12	812.5	35.9	393	1 GC1M_MOUSE	P01869 mus musculus
13	804.5	35.5	329	1 GCC_RAT	P20762 rattus norv
14	804	35.5	330	1 GCAA_MOUSE	P01863 mus musculus
15	804	35.5	399	1 GCAM_MOUSE	P01865 mus musculus
16	797	35.2	335	1 GCAB_MOUSE	P01864 mus musculus
17	780.5	34.5	322	1 GCA_RAT	P20760 rattus norv
18	774	34.2	336	1 GCB_MOUSE	P01866 mus musculus
19	774	34.2	405	1 GCBM_MOUSE	P01867 mus musculus
20	360.5	15.9	391	1 MUCB_HUMAN	P04220 homo sapien
21	353.5	15.6	421	1 EPC_MOUSE	P06336 mus musculus
22	353	15.6	429	1 EPC_RAT	P01855 rattus norv
23	353	15.6	454	1 MUC_HUMAN	P01871 homo sapien
24	352.5	15.6	455	1 MUC_MOUSE	P01872 mus musculus
25	347.5	15.3	476	1 MUCM_MOUSE	P01873 mus musculus
26	343	15.2	428	1 EPC_HUMAN	P01854 homo sapien
27	339.5	15.0	461	1 TNR2_HUMAN	P20333 homo sapien
28	337	14.9	458	1 MUC_RABIT	P03988 oryctolagus
29	332	14.7	479	1 MUCM_RABIT	P04221 oryctolagus
30	331.5	14.6	454	1 MUC_MESAU	P06337 mesocricetu
31	330.5	14.6	457	1 MUC_SUNMU	P20768 suncus mur
32	330.5	14.6	474	1 TNR2_MOUSE	P25119 mus musculus
33	329.5	14.6	450	1 MUC_CANFA	P01874 canis famil

ALIGNMENTS

RESULT 1	GC1_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG GAMMA-1 CHAIN C REGION.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
RN	[3]			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
RN	[4]			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X. Intrachain disulfide bonds."			

P23085 heterodontu
P23087 heterodontu
P01879 oryctolagus
P23088 heterodontu
P23086 heterodontu
P27512 mus musculus
P01875 gallus gall
P23084 heterodontu
P01877 homo sapien
P01876 homo sapien
P20758 gorilla gor
P25942 homo sapien

34 304.5 13.4 438 1 HVC2_HETER
35 298 13.2 438 1 HVCS_HETER
36 295 13.0 299 1 ALC_RABIT
37 294 13.0 461 1 HVCM_HETER
38 287 12.7 393 1 HVC3_HETER
39 284 12.5 289 1 CD40_MOUSE
40 282.5 12.5 446 1 MUC_CHICK
41 277.5 12.3 370 1 HVC1_HETER
42 269.5 11.9 340 1 ALC2_HUMAN
43 268.5 11.9 353 1 ALC1_HUMAN
44 267 11.8 353 1 ALC1_GORGO
45 264 11.7 277 1 CD40_HUMAN

Biochemistry 9:3188-3196(1970).
[7]
RN DISULFIDE BONDS.
RX MEDLINE-77070267; PubMed-1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT "Rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
RT characterization of the protein, the L- and H-chains, the
RT cyanogen bromide cleavage products, and the disulfide bridges.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
[8]
RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE-81208100; PubMed-7236608;
RA Deisenhofer J.;
RT "Crystallographic refinement and atomic models of a human Fc fragment
RT and its complex with fragment B of protein A from Staphylococcus
RT aureus at 2.9- and 2.8-A resolution.";
RL Biochemistry 20:2361-2370(1981).
CC GIM(1) MARKERS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE
CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)
CC MARKER & THE GIM (NON-1) MARKERS.
CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
CC 35,116,198,269 & 272.
CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
CC 268-272.
CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
CC RESIDUES 198,267&272.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00228; AAC82527.1; ALT_INIT.
DR PIR; A02146; GHU.
DR PDB; 1FC1; 15-JUL-92.
DR PDB; 1FC2; 15-JUL-92.
DR MIN; 147100;
DR INTERPRO; IPR000495;
DR INTERPRO; IPR003006;
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
3D-structure.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 223 CH2.
FT DOMAIN 224 330 CH3.
FT DISULFID 27 83
FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT CARBOHYD 180 180
FT VARIANT 97 97
FT VARIANT 239 239
FT VARIANT 241 241
FT MOD_RES 330 330
FT STRAND 123 126
FT HELIX 130 134
FT TURN 136 137
FT STRAND 141 148
FT STRAND 158 162
FT TURN 163 164

FT STRAND 165 166
FT STRAND 175 178
FT STRAND 183 190
FT HELIX 193 197
FT TURN 198 199
FT STRAND 202 206
FT STRAND 215 219
FT STRAND 227 227
FT STRAND 230 234
FT HELIX 238 240
FT TURN 241 242
FT STRAND 245 256
FT STRAND 260 266
FT TURN 267 268
FT STRAND 269 270
FT STRAND 274 276
FT STRAND 280 281
FT TURN 283 284
FT STRAND 287 296
FT HELIX 297 301
FT TURN 302 303
FT STRAND 306 312
FT TURN 313 314
FT TURN 316 317
FT STRAND 320 324
SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;

Query Match 53.2%; Score 1204; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 6.3e-74;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 TCPCPAPELLGGPSVFLPDKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 242
Db 108 TCPCPAPELLGGPSVFLPDKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 167
QY 243 HNAKTKPREQYNSTYRVSVLTVLDHQLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 302
Db 168 HNAKTKPREQYNSTYRVSVLTVLDHQLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 227
QY 303 EPQVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEESNGQPENNYKTPPVLDSDGSF 362
Db 228 EPQVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEESNGQPENNYKTPPVLDSDGSF 287
QY 363 FLYSKLTVDKSRWQQGNVSCVMHEALHNYHTOKSLSPG 404
Db 288 FLYSKLTVDKSRWQQGNVSCVMHEALHNYHTOKSLSPG 329

RESULT 2
GC2_HUMAN
ID GC2_HUMAN STANDARD; PRT; 326 AA.
AC P01859;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2 CHAIN C REGION.
GN IGHG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-82197621; PubMed-6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
RT heavy chain constant region genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
RN [2]
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
RX MEDLINE-81007873; PubMed-6774012;
RA Wang A.-C., Tung E., Fudenberg H.H.;
RT "The primary structure of a human IgG2 heavy chain: genetic,

evolutionary, and functional implications.";
 J. Immunol. 125:1048-1054(1980).
 [3]
 RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
 RX MEDLINE=80001357; PubMed=113060;
 RA Connell G.E., Parr D.M., Hofmann T.;
 RT "The amino acid sequences of the three heavy chain constant region
 RT domains of a human IgG2 myeloma protein.";
 Can. J. Biochem. 57:758-767(1979).
 [4]
 RP SEQUENCE OF 238-275 (ZIE).
 RX MEDLINE=80114419; PubMed=118920;
 RA Hofmann T., Parr D.M.;
 RT "A note of the amino acid sequence of residues 381-391 of human
 RT immunoglobulin gamma chains.";
 Mol. Immunol. 16:923-925(1979).
 [5]
 RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).
 RA Hofmann T., Parr D.M.;
 RL Submitted (MAR-1980) to the PIR data bank.
 [6]
 RP SEQUENCE OF 1-121 (DOT).
 RX MEDLINE=95255298; PubMed=7737190;
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
 RT "Characterization of the two unique human anti-flavin monoclonal
 RT immunoglobulins.";
 Eur. J. Biochem. 228:886-893(1995).
 [7]
 RP DISULFIDE BONDS.
 RX MEDLINE=72033500; PubMed=4940472;
 RA Milstein C., Frangione B.;
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
 Biochem. J. 121:217-225(1971).
 [8]
 RP DISULFIDE BONDS.
 RX MEDLINE=69064124; PubMed=5782707;
 RA Frangione B., Milstein C., Pink J.R.L.;
 RT "Structural studies of immunoglobulin G.";
 Nature 221:145-148(1969).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; V00554; CAA23814.1; -;
 DR EMBL; V00554; CAA23815.1; -;
 DR EMBL; V00554; CAA23816.1; -;
 DR EMBL; V00554; CAA23817.1; -;
 DR PIR; A02148; G2HU.
 DR MIM; 147110; -;
 DR INTERPRO; IPR000495; -;
 DR INTERPRO; IPR003006; -;
 DR PFAM; PF00047; ig; 3.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 98 CH1.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 219 CH2.
 FT DOMAIN 220 326 CH3.
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 140 200
 FT DISULFID 246 304
 FT VARIANT 60 60
 S -> A (IN MYELOMA PROTEINS TIL & ZIE).

FT SITE 156 156 /FTID=VAR_003889.
 FT MOD_RES 326 326 AT OR NEAR THE COMPLEMENT-BINDING SITE.
 FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;
 SQ
 Query Match 50.2%; Score 1137.5; DB 1; Length 326;
 Best Local Similarity 94.1%; Pred. No. 1.7e-69;
 Matches 208; Conservative 8; Mismatches 4; Indels 1; Gaps 1;
 QY 184 CPPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPPVKENWYDGVVEVH 243
 DB 106 CPPCPAPP-VAGSPVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVQFNWYDGVVEVH 164
 QY 244 NAKTKPREQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPTEKTIKAKGPRE 303
 DB 165 NAKTKPREQFNSTFRVSVLTVVHODWLNKKEYCKVSNKGLPAPTEKTIKTKGPRE 224
 QY 304 PQVYTLPPSRDELTKNOVSLTCLVKGFPSPDI AVESNGQPNENNYKTPPVLDSGSGFF 363
 DB 225 PQVYTLPPSRDEMTKNOVSLTCLVKGFPSPDI AVESNGQPNENNYKTPPVLDSGSGFF 284
 QY 364 LYSKLVKDSRWQGNVFCSCVMHEALHNHYTKSLSLSPG 404
 DB 285 LYSKLVKDSRWQGNVFCSCVMHEALHNHYTKSLSLSPG 325
 RESULT 3
 GC4_HUMAN
 ID GC4_HUMAN STANDARD; PRT; 327 AA.
 AC P01861;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-4 CHAIN C REGION.
 GN IGHG4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83157104; PubMed=6299662;
 RA Ellison J.W., Buxbaum J.N., Hood L.E.;
 RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
 RL DNA 1:11-18(1981).
 RN [2]
 RP SEQUENCE OF 1-30 AND 81-326.
 RX MEDLINE=70207560; PubMed=4192699;
 RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
 RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
 RT constant region of a gamma 4 chain.";
 RL Biochem. J. 117:33-47(1970).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; K01316; AAB59394.1; ALT_INIT.
 DR PIR; A02150; G4HU.
 DR MIM; 147130; -;
 DR INTERPRO; IPR000495; -;
 DR INTERPRO; IPR003006; -;
 DR PFAM; PF00047; ig; 3.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 98 CH1.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 220 CH2.

FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3BDBD811EF208E7A CRC64;

Query Match 50.08; Score 1133; DB 1; Length 327;
Best Local Similarity 92.98; Pred. No. 3.5e-69;
Matches 208; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 181 GGTCPCCPAPELGSPVFLFPKPKDFLMSRTPEVTCVVVDVSHEDPEVKENYVDGV 240
Db 103 GPPCSPAPFELGSPVFLFPKPKDFLMSRTPEVTCVVVDVSHEDPEVKENYVDGV 162
QY 241 EVHNAKTPREEQNSTYRVSVLTVLHODWLNGKEYCKVSKNRPALPIETISKAKGQ 300
Db 163 EVHNAKTPREEQNSTYRVSVLTVLHODWLNGKEYCKVSKNRPALPIETISKAKGQ 222
QY 301 PREQVYTLPSRDELTKQVSLCLVKGFPSPDIQVWESNGPENNKTTPPVLDSDG 360
Db 223 PREQVYTLPSRDELTKQVSLCLVKGFPSPDIQVWESNGPENNKTTPPVLDSDG 282
QY 361 SFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPG 404
Db 283 SFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPG 326

RESULT 4

GC3_HUMAN
ID GC3_HUMAN STANDARD; PRT; 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).
GN IGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
SEQUENCE (DISEASE PROTEIN WIS).
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RA MEDLINE-81021548; PubMed-6774747;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
RL gamma 3 heavy-chain disease protein Wis.";
RL Biochemistry 19:4304-4308(1980).
[2]
REVIEWS TO 12-97 OF PROTEIN WIS.
RA MEDLINE-77118561; PubMed-402363;
RA Michaelson T.E., Frangione B., Franklin E.C.;
RT "Primary structure of the 'hinge' region of human IgG3. Probable
RT quadruplication of a 15-amino acid residue basic unit.";
RL J. Biol. Chem. 252:883-889(1977).
[3]
REVIEWS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).
RX MEDLINE-77021516; PubMed-823945;
RX Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
RT Structure of the FC fragment of immunoglobulin G3.";
RL Biochem. Biophys. Res. Commun. 71:907-914(1976).
[4]
SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
RX MEDLINE-82247835; PubMed-6808505;
RX Alexander A., Steimetz M., Barriault D., Frangione B.,
RA Franklin E.C., Hood L., Buxbaum J.N.;
RT "Gamma heavy chain disease in man: cDNA sequence supports partial
RT gene deletion model.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA

CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
CC NORMALLY PRESENT IN THE HINGE REGION.
CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
CC REF.2.
CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
CC AND ALL OF THE CH1 REGION.
CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
CC GAMMA-3 HEAVY CHAINS.
CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
CC OR ANOTHER GAMMA CHAIN SUBCLASS.
CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
CC SEGMENT (12-28).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00231; AAA52805.1; ALT_SEQ.
CC PIR: A02149; G3HUI.
CC MIM: 147120;
CC INTERPRO: IPR000495; -;
CC INTERPRO: IPR003006; -;
CC PFAM: PF00047; Ig; 2.
CC PROSITE: PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT DOMAIN 12 73
FT DOMAIN 74 183
FT DISULFID 24 24
FT DISULFID 27 27
FT DISULFID 33 33
FT DISULFID 39 39
FT DISULFID 42 42
FT DISULFID 48 48
FT DISULFID 54 54
FT DISULFID 57 57
FT DISULFID 63 63
FT DISULFID 69 69
FT DISULFID 72 72
FT CARBOHYD 140 140
FT MOD_RES 290 290
FT VARIANT 126 127
FT VARIANT 134 134
FT VARIANT 139 139
FT VARIANT 182 182
FT VARIANT 227 227
FT VARIANT 227 227
FT VARIANT 279 279
FT SEQUENCE 290 AA; 32331 MW; 569CB95705B2F46 CRC64;
N-LINKED (GLCNAC...)
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
INTERCHAIN (WITH HEAVY CHAIN DIMER).
N-LINKED (GLCNAC...)
REMOVED POST-TRANSLATIONALLY.
QV -> EB (IN ZUC).
/FTid=VAR_003890.
P -> L (IN OMM).
/FTid=VAR_003891.
F -> Y (IN OMM).
/FTid=VAR_003892.
T -> A (IN OMM).
/FTid=VAR_003893.
S -> N (IN OMM).
/FTid=VAR_003894.
MISSING (IN ZUC).
/FTid=VAR_003895.
F -> Y (IN OMM).
/FTid=VAR_003896.
569CB95705B2F46 CRC64;

Query Match

49.7%; Score 1125.5; DB 1; Length 290;

Best Local Similarity 69.6%; Pred. No. 9.8e-69;
Matches 217; Conservative 19; Mismatches 23; Indels 53; Gaps 5;

QY 97 CLAHRSO---PPGFGVQAGTPERNITCKPCPDGFFSNETSSKAP-CRKHTNCSVFGLLL 152
Db 27 CPKPKSCDTPPP-----CPKCPK-KSCDTPPPCPKCPKPKSCDT----- 65

QY 153 TQKGNATHDNICSGNSESTQKSGGGGGGTGCPAPPELLGGPSVFLPPKPKDILMIS 212
Db 66 -----PPCPCPCAPPELLGGPSVFLPPKPKDILMIS 97

QY 213 RTPVTCVVVDVSHEDPEVKFNKYVDGVEVHNNAKTPREOYNSTYRVSVLTVLHQDWL 272
Db 98 RTPVTCVVVDVSHEDPEVQFKYVDGVQVHNNAKTPREOQFNSTFRVSVLTVLHQDWL 157

QY 273 NGKEYCKVSNKALPAPIETKISKAKGPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 332
Db 158 DGKEYCKVSNKALPAPIETKISKAKGPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 217

QY 333 SDIAVESNGQPNENYKTPPPVLDSDGSPFLYSKLTVDKSRWQOGNVFSCSVMEALHN 392
Db 218 SDIAVESNGQPNENYKTPPPVLDSDGSPFLYSKLTVDKSRWQOGNVFSCSVMEALHN 277

QY 393 HYTKQSLSLSPG 404
Db 278 RFTQKSLSLSPG 289

RESULT 5

GC_RABIT STANDARD; PRT; 323 AA.

ID GC_RABIT
AC P01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 01, Last sequence update)
DE IG GAMMA CHAIN C REGION.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
F-1 haplotype."
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
immunoglobulin G of different allotype."
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steimetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: Isolation of a cDNA encoding gamma
heavy chain and identification of two genomic C gamma genes."
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
immunoglobulin G."
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (in) Killefer J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
Stockholm (1967).

CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER, E15
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M16426; AAA31289.1; -
DR PIR: A02161; GHRB.
DR INTERPRO: IPR000495; -
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> E (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).
FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 E -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 40.2%; Score 909; DB 1; Length 323;
Best Local Similarity 72.8%; Pred. No. 3.3e-54;
Matches 163; Conservative 28; Mismatches 31; Indels 2; Gaps 1;

QY 183 TC--PPCAPPELLGGPSVFLPPKPKDILMISRTPEVTCVVVDVSHEDPEVKENYVDGV 240
Db 99 TCSKPTCPPELLGGPSVFLPPKPKDILMISRTPEVTCVVVDVSHEDPEVKENYVDGV 158

QY 241 EVHNAKTPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVKYKALPAPIETKISKARGQ 300
Db 159 QVTRAPPLREQQPNSTIRVSVLTPLTHQDWLNGKEYCKVKYKALPAPIETKISKARGQ 218

QY 301 PREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAVESNGQPNENYKTPPPVLDSDG 360
Db 219 PLEPKVYTMGPPEELSLRSVSLTCMNGFYPSDISVSEWEKNGKAEDNYKTPPAVLSDG 278

QY 361 SEFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTKQSLSLSPG 404
Db 279 SYFLYNKSLVPTSEWQRGDVFSCSVMEALHNHYTKQSLSLSPG 322

RESULT 6

GC2_CAVPO STANDARD; PRT; 329 AA.

ID GC2_CAVPO
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2 CHAIN C REGION
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN [1]
RP SEQUENCE OF 1-3.

DE IG GAMMA-1 CHAIN C REGION.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89232738; PubMed=3149946;
 RA Brueggemann M.;
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
 RL Gene 74:473-482(1988).
 DR PIR: PS0017; PS0017.
 DR INTERPRO: IPR000495; -
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; ig; 3.
 DR PROSITE: PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
 FT NON_TER 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 112 HINGE.
 FT DOMAIN 113 219 CH2.
 FT DOMAIN 220 326 CH3.
 FT DISULFID 27 82
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 140 200
 FT DISULFID 246 304
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;
 Query Match 36.48; Score 824.5; DB 1; Length 326;
 Best Local Similarity 63.08; Pred. No. 1.5e-48;
 Matches 143; Conservative 41; Mismatches 36; Indels 7; Gaps 2;
 Qy 181 GGTGPPCPAPPELLGG---PSVLEPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 237
 Db 103 GGDCKPC---ICTGSEVSSVFIFPPPKDVLITLTPKVCVVVDISQDDPEVHFSWFV 158
 Qy 238 DGVEVHNAKTPRQEQNSTRVSVLTFLVDQLNGKEYCKVSKNKPALPAPIEKTISKA 297
 Db 159 DDVEVHTAQTTPPEQNFSTRSVSELPILHQLDNLNGRTFRCVKVTSAAFPSPIEKTIKSP 218
 Qy 298 KGQPREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVEGSGNQENPKYKTPPVLD 357
 Db 219 EGRTQVPHVYTSPTKEKNTQVSIICMVKGFYPPDIIYVQWNGQEQENYKNTPTMD 278
 Qy 358 SDGSFFLYSLTKVTKSRWQQGNVFSCSYMHEALNHNHYTKSLSPG 404
 Db 279 TDGSVFLYSLKLVKKEKWKQGNFTFCSVLHREGLHNHHTPEKSLSHSPG 325
 RESULT 11
 ID GCL_MOUSE STANDARD; PRT; 324 AA.
 AC P01868;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE IG GAMMA-1 CHAIN C REGION.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 gamma 1 chain gene.";
 RL Cell 18:559-568(1979).
 RN [2]

RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
 RX MEDLINE=80202559; PubMed=6769752;
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
 cloned in a bacterial plasmid.";
 RL Gene 9:87-97(1980).
 RN [3]
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=80012837; PubMed=113776;
 RA Rogers J., Clarke P., Salsner W.;
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
 heavy chain.";
 RL Nucleic Acids Res. 6:3305-3321(1979).
 RN [4]
 RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=78242288; PubMed=98524;
 RA Adetugbo K.;
 RT "Evolution of immunoglobulin subclasses. Primary structure of a
 murine myeloma gammal chain.";
 RL J. Biol. Chem. 253:6068-6075(1978).
 RN [5]
 RP DISULFIDE BONDS (MOPC 21).
 RX MEDLINE=73008889; PubMed=5073237;
 RA Svasti J., Milstein C.;
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
 RL Biochem. J. 126:837-850(1972).
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00793; CAA24172.1; -
 CC EMBL; V00793; CAA24173.1; -
 CC EMBL; V00793; CAA24174.1; -
 CC EMBL; V00793; CAA24175.1; -
 CC EMBL; V00795; CAA24176.1; -
 CC PIR; A02159; GIMS.
 CC MGD; MGI:96446; Igh-4.
 CC INTERPRO: IPR000495; -
 CC INTERPRO: IPR003006; -
 CC PFAM: PF00047; ig; 3.
 CC PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW Alternative splicing.
 FT NON_TER 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 110 HINGE.
 FT DOMAIN 111 217 CH2.
 FT DOMAIN 218 324 CH3.
 FT DISULFID 27 82
 FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 138 198
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
 FT DISULFID 244 302
 FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
 FT CONFLICT 276 276 N -> D (IN REF. 3).
 FT CONFLICT 278 278 N -> D (IN REF. 3).
 SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;
 Query Match 35.98; Score 812.5; DB 1; Length 324;
 Best Local Similarity 62.38; Pred. No. 9.4e-48;
 Matches 139; Conservative 44; Mismatches 35; Indels 5; Gaps 2;
 Qy 184 CPP--CPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDCGE 241

```

Db 104 CKPCICTVPEV---SSVFIPPPKDKVLTITLTPKVCVVVDISKDDPEVQFSWFVDVDE 160
QY 242 VHNAKTKPREOYNSTYRVVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTISKAKGP 301
Db 161 VHTAQTPREOQNSFRSVSELPIMHQDWLNGKEFKCRVNSAFAFPAPIEKTISKTKGRP 220
QY 302 REPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGQPENNYKTTPPVLDSDGS 361
Db 221 KAPOVVTIPPPKQMAKDKVSLTCMTDFEPDITVWQNGQPAENYKNTQPIMTNGS 280
QY 362 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPG 404
Db 281 YFYVSLNVLQVKSWEAGNFTCSVLHGLHNHHTKSLSHSPG 323

RESULT 12
GCCIM_MOUSE STANDARD; PRT; 393 AA.
ID GCCIM_MOUSE AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=82221190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
SEGMENT OF MU CHAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC

```

```

DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02155; GLMSW.
DR MGD; MGI:96446; IGH-4.
DR INTERPRO; IPR000495; -
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; I9; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 138
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393
FT SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;
QY 184 CPP--CPAPELLGGPSVFLPPPKDPTLMISRTPETVTCVVVDVSHEDPEVKFNNYVDGVE 241
Db 104 CKPCICTVPEV---SSVFIPPPKDKVLTITLTPKVCVVVDISKDDPEVQFSWFVDVDE 160
QY 242 VHNAKTKPREOYNSTYRVVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTISKAKGP 301
Db 161 VHTAQTPREOQNSFRSVSELPIMHQDWLNGKEFKCRVNSAFAFPAPIEKTISKTKGRP 220
QY 302 REPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGQPENNYKTTPPVLDSDGS 361
Db 221 KAPOVVTIPPPKQMAKDKVSLTCMTDFEPDITVWQNGQPAENYKNTQPIMTNGS 280
QY 362 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPG 404
Db 281 YFYVSLNVLQVKSWEAGNFTCSVLHGLHNHHTKSLSHSPG 323

RESULT 13
GCC_RAT
ID GCC_RAT AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2C CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
region cDNA: extensive homology to mouse gamma 3.";
RL Eur. J. Immunol. 18:317-319(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

```


ID GCAM_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-8222190; PubMed-6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; J00471; AAB59661.1; ALT_INIT.
DR PIR; A02154; G2MSAM.
DR MGD; MGI:96443; IGH-1.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR003006; -.
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 15 15
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FE0 CRC64;

Query Match 35.5%; Score 804; DB 1; Length 399;
Best Local Similarity 65.0%; Pred. No. 4.4e-47;
Matches 145; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

QY 184 CPP--CPAPELLGGPSVFLPPPKPTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVE 241
DB 107 CPPCKCPAPMLLGGPSVFIPPPKIKDVLMSLSPITVCVVVDVSEDDPDVQISFVNNVE 166
QY 242 VHNAKTKPREEQYNSTYRVSVLTVLHQDWLNGKEYKCKVSKNALKALPAPIETISKAKGQP 301
DB 167 VHTAQITQTHREDYNSLTIRVVSALPIQHDWMSGKEFKCKVNNKDLPAPIETISKPKGSV 226
QY 302 REPOVYTLPPSKNQLKQVSLCLVKGYFSPDAVWEESNQGPENNKKTPPVLDSGDS 361
DB 227 RAPQVYVLPPEEMTKKQVTLTCMVTDMPEDIIYVWNTNNGKTELNYKNKTEPVLDSGDS 286

QY 362 FFYLSKLTVDKSRWQGNVFSQSVMHSEALHNHYTKSLSPG 404
DB 287 YFMYSKLRVEKKNNWVERNSYSCSVVHEGLNHHHTTKSFSTPG 329

Search completed: March 1, 2001, 09:17:48
Job time: 266 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:48 ; Search time 79.26 Seconds
(without alignments)
346.100 Million cell updates/sec

Title: US-09-389-782A-7
Perfect score: 2264
Sequence: 1 ETFFPKYLHYDEETSHQLLC.....VMHEALHNHYTKQSLSLSPG 404

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1204	53.2	255	4 S31866	Ig gamma-1 chain C
2	1204	53.2	330	1 GHUU	Ig heavy chain V r
3	1202	53.1	374	2 S72664	Ig heavy chain V r
4	1198	52.9	374	2 S69339	Ig gamma chain C r
5	1162	51.3	234	2 PT0207	Ig gamma-3 chain C
6	1146.5	50.6	377	2 A23511	Ig gamma-3 chain C
7	1144.5	50.6	377	2 A60764	Ig gamma-2 chain C
8	1137.5	50.2	326	1 G2HU	Ig gamma-4 chain C
9	1133	50.0	327	1 G4HU	Ig gamma-3 heavy c
10	1125.5	49.7	289	1 G3HUW1	Ig gamma chain C r
11	909	40.2	323	1 GHRB	Ig gamma 2b chain
12	902.5	39.9	328	2 I47160	Ig gamma 2a chain
13	902.5	39.9	328	2 I47159	Ig gamma 4 chain c
14	900.5	39.8	277	2 I47162	Ig gamma-2 chain c
15	890.5	39.3	329	1 G2GP	Ig gamma-1 chain c
16	880.5	38.9	328	2 I47158	Ig gamma 1 chain c
17	874.5	38.6	328	2 I47161	Ig gamma 3 chain c
18	850.5	37.6	470	2 S22080	Ig heavy chain pre
19	841.5	37.2	329	1 G3MSC	Ig gamma-3 chain C
20	841	37.1	308	2 C30554	Ig heavy chain C r
21	841	37.1	472	2 S31459	Ig gamma-1 chain -
22	840	37.1	333	2 PS0018	Ig gamma-2b chain
23	835.5	36.9	398	1 G3MSM	Ig gamma-3 chain C
24	824.5	36.4	326	2 PS0017	Ig gamma-1 chain C
25	822.5	36.3	444	2 PC4136	monoclonal antibody
26	812.5	35.9	324	1 G1MS	Ig gamma-1 chain C
27	812.5	35.9	393	1 G1MSM	Ig gamma-1 chain C
28	804.5	35.5	329	2 S00847	Ig gamma-2c chain
29	804	35.5	330	1 G2MSA	Ig gamma-2a chain

30	804	35.5	399	1 G2MSAM	Ig gamma-2a chain
31	804	35.5	469	2 S37483	Ig gamma-2a chain
32	797	35.2	335	1 G2MSAB	Ig gamma-2a chain
33	794	35.1	446	2 S40295	Ig gamma-2a chain
34	780.5	34.5	322	2 PS0019	Ig gamma-2a chain
35	774	34.2	336	1 G2MS11	Ig gamma-2b chain
36	774	34.2	405	1 G2MSBM	Ig gamma-2b chain
37	774	34.2	474	2 S25057	Ig gamma-2b chain
38	760.5	33.6	327	2 S06611	Ig gamma-2 chain C
39	752	33.2	475	2 S01321	Ig gamma-2b chain
40	702	31.0	180	2 I46732	Ig gamma heavy cha
41	577.5	25.5	249	2 S69340	Ig heavy chain VHI
42	574.5	25.4	218	2 A36040	Ig heavy chain V-I
43	566	25.0	152	2 S14236	Ig gamma-1 chain C
44	402	17.8	572	2 B46529	Ig Y heavy chain (
45	361	15.9	549	2 S04845	Ig Y heavy chain pre

ALIGNMENTS

RESULT 1

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C:Accession: S31866

R:Filpula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene produc

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FIL>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 53.2%; Score 1204; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 3e-68;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 TCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 242

Db 33 TCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 92

QY 243 HNAKTKPREQYNSTYRVVSVLTVLHODWLNQKEYCKVSNKALPAPTEKTSKAKGQPR 30

Db 93 HNAKTKPREQYNSTYRVVSVLTVLHODWLNQKEYCKVSNKALPAPTEKTSKAKGQPR 15

QY 303 EPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVSNQGPENNYKTTTPVLDSDGSF 362

Db 153 EPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVSNQGPENNYKTTTPVLDSDGSF 212

QY 363 FLYSKLTVDKSRWQOGNWFSCVMHEALHNHYTKQSLSPG 404

Db 213 FLYSKLTVDKSRWQOGNWFSCVMHEALHNHYTKQSLSPG 254

RESULT 2

GHUU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A30564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma gene.

A:Reference number: A93433; MUID:82274238

A:Accession: A93433

A:Molecule type: DNA
A:Residues: 1-330 <ELL>
A:Cross-references: EMBL:Z17370
A:Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,
A:Note: Lys-330 is removed after translation
R:Harris, L.J.
submitted to the EMBL Data Library, October 1992
A:Reference number: S33904
A:Accession: S36861
A:Molecule type: DNA
A:Residues: 2-330 <HAR>
A:Cross-references: EMBL:Z17370
R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A:Title: Structure of human immunoglobulin gamma genes: Implications for evolution of a
A:Reference number: S33887; MUID:83001943
A:Accession: S33887
A:Molecule type: DNA
A:Residues: 88-113:235-330 <TAK>
A:Cross-references: EMBL:Z17370
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,
Biochemistry 9, 3161-3170, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen
A:Reference number: A90563; MUID:71064024
A:Contents: myeloma protein Eu
A:Accession: B90563
A:Molecule type: protein
A:Residues: 1-96,'R',98-135 <CUN>
A:Note: this sequence has the G1m(3) marker, 97-Arg
R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequen
A:Reference number: A90564; MUID:71064025
A:Contents: Eu
A:Accession: A90564
A:Molecule type: protein
A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',240,
A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met
R:Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie),
igen Primärstruktur.
A:Reference number: A91668; MUID:77070269
A:Contents: myeloma protein Nie
A:Accession: B91668
A:Molecule type: protein
A:Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27
A:Note: this sequence has the G1m(17) and G1m(1) markers
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KO
A:Reference number: A91723; MUID:83289131
A:Contents: myeloma protein KO; disulfide bonds
A:Accession: A91723
A:Molecule type: protein
A:Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH
A:Note: this sequence has the G1m(3) and G1m(non-1) markers
R:Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
A:Reference number: A90565; MUID:71064027
A:Contents: annotation; disulfide bonds
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A:Reference number: A91667; MUID:77070267
A:Contents: annotation; disulfide bonds
C:Genetics:
A:Gene: GDB:IGH1
A:Cross-references: GDB:120085; OMTM:147100
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 114/1; 224/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/Disulfide bonds: #status experimental
F:103/Disulfide bonds: interchain (to light chain) #status experimental
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 53.2%; Score 1204; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 4e-68;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 183 TCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 242
Db 108 TCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 167

Qy 243 HNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPR 302
Db 168 HNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPR 227

Qy 303 EPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF 362
Db 228 EPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF 287

Qy 363 FLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPG 404
Db 288 FLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPG 329

RESULT 3
S72664
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
C:Accession: S72664
R:Khamilichi, A.A.
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 53.18%; Score 1202; DB 2; Length 374;
Best Local Similarity 90.8%; Pred. No. 6e-68;
Matches 226; Conservative 5; Mismatches 10; Indels 8; Gaps 2;

Qy 156 GNATHDNICNSBSTOKSGGGGGGCPAPPELLGGPSVFLFPPPKDRLMISRT 215
Db 133 GQGLTLTVCS-EPKSCDKT-----HTCPAPPELLGGPSVFLFPPPKDRLMISRT 184

Qy 216 EVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGK 275
Db 185 EVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGK 244

Qy 276 EYCKVSNKALPAPIEKTISKAKGQPRQVYTLPPSRDELTKNOVSLTCLVKGFYPSDI 335
Db 245 EYCKVSNKALPAPIEKTISKAKGQPRQVYTLPPSRDELTKNOVSLTCLVKGFYPSDI 304

Qy 336 AVEWESNGQPENNYKTTPPVLDSDGSFELYSKLTVDKSRWQGNVFCSSVMHEALHNHYT 395
Db 305 AVEWESNGQPENNYKTTPPVLDSDGSFELYSKLTVDKSRWQGNVFCSSVMHEALHNHYT 364

Qy 396 QKSLSLSPG 404
|||||

Db 365 QKSLSPG 373

RESULT 4

Ig heavy chain V region precursor - human

C:Species: Homo sapiens (man)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-2000

C:Accession: S69339

R:Khamichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.

Eur. J. Biochem. 229, 54-60, 1995

A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.

A:Reference number: S69339; MUID:95262687

A:Accession: S69339

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <KHA>

A:Cross-references: EMBL:X81695

C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 52.9%; Score 1198; DB 2; Length 374;
Best Local Similarity 99.1%; Pred. No. 1.1e-67;
Matches 220; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 183 TCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEV 242

Db 152 TCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEV 211

QY 243 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 302

Db 212 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 271

QY 303 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSF 362

Db 272 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSF 331

QY 363 FLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTQKSLSPG 404

Db 332 FLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTQKSLSPG 373

RESULT 5

PT0207

Ig gamma chain C region - chimpanzee

C:Species: Pan troglodytes (chimpanzee)

C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999

C:Accession: PT0207

R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.

Mol. Immunol. 28, 319-322, 1991

A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.

A:Reference number: PT0207; MUID:91287716

A:Accession: PT0207

A:Molecule type: mRNA

A:Residues: 1-234 <EHR>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 51.3%; Score 1162; DB 2; Length 234;
Best Local Similarity 99.1%; Pred. No. 1.1e-65;
Matches 214; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 183 TCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEV 242

Db 19 TCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEV 78

QY 243 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 302

Db 79 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 138

QY 303 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSF 362

Db 139 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSF 198

QY 363 FLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTQKS 398

Db 199 FLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTQKS 234

RESULT 6

A23511

Ig gamma-3 chain C region (allotype G3m(b)) - human

C:Species: Homo sapiens (man)

C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999

C:Accession: A23511

R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.

Nucleic Acids Res. 14, 1779-1789, 1986

A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene.

A:Reference number: A23511; MUID:86148507

A:Accession: A23511

A:Molecule type: DNA

A:Residues: 1-377 <HUC>

A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056

C:Genetics:

A:Gene: GDB:IGHG3

A:Cross-references: GDB:119339; OMIM:147120

A:Map position: 14q32.33-14q32.33

A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 50.6%; Score 1146.5; DB 2; Length 377;
Best Local Similarity 66.0%; Pred. No. 1.7e-64;
Matches 227; Conservative 19; Mismatches 33; Indels 65; Gaps 7;

QY 76 CNRTHNRCVCECKGYLEIEF-----CLKHRSC---PPFGVVOAGTPEPNTVC 121

Db 83 CYNVH-KPSNTKVDKRVELKTLPLGDTTHTCPRCEPKSCDTPP-----C 126

QY 122 KRCPDGFNSNETSKAP-CRKHTNCVSFGLLLLTOKGNATHDNCISGNSESTQKSGGGGG 180

Db 127 PRCPPEP-KSCDTPPCPRCPKSCDT----- 152

QY 181 GGTCCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGV 240

Db 153 PPCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGV 212

QY 241 EVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ 300

Db 213 EVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ 272

QY 301 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDG 360

Db 273 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDG 332

QY 361 SFFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTQKSLSPG 404

Db 333 SFFLYSKLTVDKSRWQOGNVFSCSVNHEALHNNHYTQKSLSPG 376

RESULT 7

A60764

Ig gamma-3 chain C region, form LAT - human

C:Species: Homo sapiens (man)

C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999

C:Accession: A60764

R:Huck, S.; Lefranc, G.; Lefranc, M.P.

Immunogenetics 30, 250-257, 1989

A:Title: A human immunoglobulin IGHG3 allele (Gmb0, bl, c3, c5, u) with an IGHG4 conv

A:Reference number: A60764; MUID:90007613

A:Accession: A60764

A>Status: preliminary

Biochem. J. 117, 33-47, 1970
A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant
A:Reference number: A90249; MUID:70207560
A:Accession: A90249
A:Molecule type: protein
A:Residues: 1-30;81-326 <PIN>
C:Genetics:
A:Gene: GDB:IGHG4
A:Cross-references: GDB:119340; OMIM:147130
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 111/1; 221/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) heavy disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds. Interchain (to light chain) #status experimental
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMI>
F:99-110/Region: hinge
F:134-203/Domain: immunoglobulin homology <IM2>
F:240-307/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:27-83,141-201,247-305/Disulfide bonds: #status predicted
F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.0%; Score 1133; DB 1; Length 327;
Best Local Similarity 92.9%; Pred. No. 1e-63; 8; Mismatches 0; Gaps 0;
Matches 208; Conservative 8; Indels 0;

QY 181 GTCPCPCAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVDVSHEDPEVKFNWYVDG 240
Db 103 GPPCPCAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVDVSHEDPEVKFNWYVDG 162

QY 241 EVHNATKPREQNSRYRVSIVLVTLHQLDNLCKEYKCKVSNKALPAPIEKTIISKAKQ 300
Db 163 EVHNATKPREQNSRYRVSIVLVTLHQLDNLCKEYKCKVSNKALPAPIEKTIISKAKQ 222

QY 301 PREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDG 360
Db 223 PREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDG 282

QY 361 SFELYSLKLVDRKQWQGVFSCSVMEALHNHYTKQSLSLSPG 404
Db 283 SFELYSLKLVDRKQWQGVFSCSVMEALHNHYTKQSLSLSPG 326

RESULT 10
G3HOWI
Ig gamma-3 heavy chain disease proteins - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C:Accession: A90442; A92219; A90198; A93915; A02149
R:Frangione, B.; Rosenwasser, E.; Prell, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-
A:Reference number: A90442; MUID:81021548
A:Contents: heavy chain disease protein Wis
A:Accession: A90442
A:Molecule type: protein
A:Residues: 1-289 <FRA>
A:Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
A:Note: the sequence of residues 42-76 was taken from the reference that follows
R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
A:Reference number: A92219; MUID:71118561
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein w
A:Accession: A92219
A:Molecule type: protein
A:Residues: 12-97 <MIC>
A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
idue segment (12-28)

A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form in
R:Wolfenstein-Todel, C.; Frangione, B.; Prell, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of t
A:Reference number: A90198; MUID:77021516
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residu
A:Accession: A90198
A:Molecule type: protein
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>
A:Note: this protein lacks most of the V region, all of the CH1 region, and part of t
R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood,
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deleti
A:Reference number: A93915; MUID:82247835
A:Contents: heavy chain disease protein Omm
A:Accession: A93915
A:Molecule type: mRNA
A:Residues: 12-70;72-114;116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-
A:Note: a carboxyl-terminal Lys is removed posttranslationally
A:Note: this sequence may represent an allelic form or another gamma chain subclass
C:Comment: The heavy chain disease protein Wis is shown.
C:Genetics:
A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
F:203-270/Domain: immunoglobulin homology <IMM>
F:1/Modified site: pyroglutamic acid (Gln) #status experimental
F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 49.7%; Score 1125, 5; DB 1; Length 289;
Best Local Similarity 69.6%; Pred. No. 2.7e-63;
Matches 217; Conservative 19; Mismatches 23; Indels 53; Gaps 5;

QY 97 CLKHRS-PPGFGVVVQAGTPERTVTCRCPCDGFNETSSKAP-CRKHTNCSVFGLL 152
Db 27 CPEPKSCDTPPP-----CPRCEP-KSCDTPPCPCPEPKSCDT----- 65

QY 153 TQGNATHNICSGNSESTOKSGGGGGGTCTPCPCAPPELLGGPSVFLFPPKPKDTLMIS 212
Db 66 -----PPCPCPCAPPELLGGPSVFLFPPKPKDTLMIS 97

QY 213 RPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQNSRYRVSIVLVTLHQLDNL 272
Db 98 RPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQNSRYRVSIVLVTLHQLDNL 157

QY 273 NGREYKCKVSNKALPAPIEKTIISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFY 332
Db 158 DGREYKCKVSNKALPAPIEKTIISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFY 217

QY 333 SDIAVESNGQPNKYKTTTPVLDSDGSPFLYSLKLVDRKQWQGVFSCSVMEALHN 392
Db 218 SDIAVESNGQPNKYKTTTPVLDSDGSPFLYSLKLVDRKQWQGVFSCSVMEALHN 277

QY 393 HYTEKSLSLSPG 404
Db 278 RFTQKSLSLSPG 289

RESULT 11
GHRB
Ig gamma chain C region - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
C:Accession: A91749; A90290; A93928; A90245; A94416; A02161
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I hsp
A:Reference number: A91749; MUID:84030930
A:Accession: A91749
A:Molecule type: mRNA

A:Residues: 1-323 <BER>
A:Note: This sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Thr
R:Pratt, D.M.; Mole, L.E.
Biochem. J. 151, 337-349, 1975
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglobulin
A:Reference number: A90290; MUID:76135469
A:Accession: A90290
A:Molecule type: protein
A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>
R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
A:Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy chain
A:Reference number: A93928; MUID:83299917
A:Accession: A93928
A:Molecule type: mRNA
A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>
A:Cross-references: GB:M16426; NID:9165111; PID:AAA31289.1; PID:9165112
A:Note: This sequence has the d12 allotypic marker, 104-Met, and the e15 allotypic marker
R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
Biochem. J. 116, 249-259, 1970
A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin
A:Reference number: A90245; MUID:70110015
A:Accession: A90245
A:Molecule type: protein
A:Residues: 132-143, 'E', 145-161 <FRU>
R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
In Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almquist and Wiksell,
A:Reference number: A94416
A:Accession: A94416
A:Molecule type: protein
A:Residues: 129-131; 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, 'Q',
A:Note: This has the e15 allotypic marker, 185-Ala
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger multimers.
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin homology
F:130-199/Domain: immunoglobulin homology <IM1>
F:20-82/Domain: immunoglobulin homology <IM2>
F:236-303/Domain: immunoglobulin homology <IM3>
F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.2%; Score 909; DB 1; Length 323;
Best Local Similarity 72.8%; Pred. No. 8.9e-50;
Matches 163; Conservative 28; Mismatches 31; Indels 2; Gaps 1;

QY 183 TC--PPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGV 240
Db 99 TCSKPTCPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGV 158

QY 241 EVHNAKTPREEQNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAKGQ 300
Db 159 QVTRAPPLRQGFNSTIRVSTVPIHQDWLNGKEYCKVSNKALPAPIETKISKAKGQ 218

QY 301 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDG 360
Db 219 PLEPKVYTMGPPEELSSRSYSLTCMINGFYPDSISVEWERKGAEDNYKTPPVLDSDG 278

QY 361 SFFLYSKLTVDKSRWQQGNVSCVMHEALHNYHTQKSLSPG 404
Db 279 SYFLYKLSVPTSEWQGRDVFCTCSVMHEALHNYHTQKSLSPG 322

RESULT 12
147160
Ig gamma 2b chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: 147160
R:Kaczkovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: 147158; MUID:95015845

A:Accession: 147160
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AAA52218.1; PID:9433126
C:Genetics:
A:Gene: IgG2b
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 39.9%; Score 902.5; DB 2; Length 328;
Best Local Similarity 59.7%; Pred. No. 2.3e-49;
Matches 172; Conservative 36; Mismatches 53; Indels 27; Gaps 4;

QY 125 PDGFFSNETSCKAPCRKHTNCSVFGLLLTQGNATHDNCISGSESTQKSGGGGGGTC 184
Db 59 PSGLYSLSSWTVTPASSLSKSY-----TCNVNHPATTTKDKRVRGTTK 103

QY 185 PPCP-----APELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 238
Db 104 PPCPICPACESP----GPSVFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 159

QY 239 GVEVHNAKTPREEQNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 298
Db 160 GVEVHTAQTREKPEQNFSTYRVSVLPIHQDWLNGKEYCKVSNKALPAPIETKISKAK 219

QY 299 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTPPV 356
Db 220 GQTPREPQVYTLPPHAELSRKSVITCLVIGYPPDIDVWQRNGQPEPEGYRTTPPQ 279

QY 357 DSDGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNYHTQKSLSPG 404
Db 280 DVDGTIFYLFYSVDKASWQGGIFQCAVMHEALHNYHTQKSLKTPG 327

RESULT 13
147159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: 147159
R:Kaczkovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: 147158; MUID:95015845
A:Accession: 147159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03779; NID:9433123; PIDN:AAA52217.1; PID:9433124
C:Genetics:
A:Gene: IgG2a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 39.9%; Score 902.5; DB 2; Length 328;
Best Local Similarity 59.7%; Pred. No. 2.3e-49;
Matches 172; Conservative 36; Mismatches 53; Indels 27; Gaps 4;

QY 125 PDGFFSNETSCKAPCRKHTNCSVFGLLLTQGNATHDNCISGSESTQKSGGGGGGTC 184
Db 59 PSGLYSLSSWTVTPASSLSKSY-----TCNVNHPATTTKDKRVRGTTK 103

QY 185 PPCP-----APELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 238
Db 104 PPCPICPACESP----GPSVFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 159

QY 239 GVEVHNAKTPREEQNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 298
Db 160 GVEVHTAQTREKPEQNFSTYRVSVLPIHQDWLNGKEYCKVSNKALPAPIETKISKAK 219

QY 299 GQPREQVYTLPPGRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTTPVL 356
Db 220 GQPREQVYTLPPHAEELSRKSVITCLVIGFYPPDIDVWQRNGQPEPEGNRYTTPQQ 279
QY 357 DSDGSFELYSLTVDKSRWQGNVFCSSVMEALHNHYTKSLSPG 404
Db 280 DVDGTGYFLYSKFSVDKASWGGGIFQCAVMEALHNHYTKSLKTPG 327

RESULT 14

I47162

Ig gamma 4 chain constant region - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000

C:Accession: I47162

R:Kacskovics, I.; Sun, J.; Butler, J.E.

J. Immunol. 153, 3565-3573, 1994

A:Title: Five putative subclones of swine IgG identified from the cDNA sequences of a

A:Reference number: I47158; MUID:95015845

A:Accession: I47162

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-277 <KAC>

A:Cross-references: EMBL:U03782; NID:g433129; PIDN:AAA52220.1; PID:g433130

C:Genetics:

A:Gene: IgG4

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 39.8%; Score 900.5; DB 2; Length 277;
Best Local Similarity 60.6%; Pred. No. 2.5e-49;
Matches 172; Conservative 35; Mismatches 58; Indels 19; Gaps 4;

QY 125 PDGFSSNETSSKAPCRKHTNCVSFGLLLTKQGNATHDNICSGNSESTQKSGGGGGGTC 184
Db 8 PSGLYLSMTVTPASSLSKSY-----TCNVNHPATTTKDKRVGKTK 52

QY 185 PCPC-APELLG-GPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 242
Db 53 PCPCIPACGPGPSAFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 112

QY 243 HNARTKPREQYNSTYRVVSVLTVHLQDNLNKGKEYCKVSKNKPAPLEKTSKAKGQPR 302
Db 113 HTAQTRPEQENSTYRVSVSLPQHODWLNKGEKCKVKNKLPAPITRIISKAKGQTR 172

QY 303 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTTPVLSDG 360
Db 173 EPQVYTLPPTEELSRKSVITCLVIGFYPPDIDVWQRNGQPEPEGNRYTTPQQDVG 232

QY 361 SFELYSLTVDKSRWQGNVFCSSVMEALHNHYTKSLSPG 404
Db 233 TYFLYSLAVDKASWQKGDITFOCAVMEALHNHYTKSLKTPG 276

RESULT 15

G2GP

Ig gamma-2 chain C region - guinea pig

C:Species: Cavia porcellus (guinea pig)

C:Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 16-Jul-1999

C:Accession: A94553; A90352; A90359; A90384; A90385; A02151

R:Trischmann, T.M.

submitted to the Atlas, April 1975

A:Reference number: A94553

A:Accession: A94553

A:Molecule type: protein

A:Residues: 1-3 <TRI>

R:Birshtein, B.K.; Hussain, Q.Z.; Cebra, J.J.

Biochemistry 10, 18-25, 1971

A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Am

A:Reference number: A90352; MUID:71058471

A:Accession: A90352

A:Molecule type: protein

A:Residues: 4-68 <BIR>
R:Turner, K.J.; Cebra, J.J.
Biochemistry 10, 9-17, 1971
A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II.
A:Reference number: A90359; MUID:71058486
A:Accession: A90359
A:Molecule type: protein
A:Residues: 69-133;312-329 <TUR>
R:Tracey, D.E.; Cebra, J.J.
Biochemistry 13, 4796-4803, 1974
A:Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibody
A:Reference number: A90384; MUID:75036072
A:Accession: A90384
A:Molecule type: protein
A:Residues: 134-226 <TRA>
R:Trischmann, T.M.; Cebra, J.J.
Biochemistry 13, 4804-4811, 1974
A:Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibody
A:Reference number: A90385; MUID:75036073
A:Accession: A90385

A:Molecule type: protein
A:Residues: 227-311 <TR2>
R:Oliveira, B.; Lamm, M.E.
Biochemistry 10, 26-31, 1971
A:Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.
A:Reference number: A90354; MUID:71058474
A:Contents: annotation; disulfide bonds
A:Note: Cys-16 is involved in a heavy-light chain bond
A:Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds
C:Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotrimer; heterotrimer; immunoglobulin F:21-81/Domain: immunoglobulin homology <IMI>
F:135-204/Domain: immunoglobulin homology <IM2>
F:241-310/Domain: immunoglobulin homology <IM3>
F:28-79/Disulfide bonds: #status experimental
F:142-202/Disulfide bonds: #status experimental
F:178/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:248-308/Disulfide bonds: #status experimental

Query Match 39.38%; Score 890.5; DB 1; Length 329;
Best Local Similarity 60.7%; Pred. No. 1.3e-48;
Matches 170; Conservative 33; Mismatches 64; Indels 13; Gaps 2;

QY 127 GFFSNETSSKAPCRKHTNCVSFGLLLTKQGNATHDNICSGNSESTQKSGGGGGGTCPP 186
Db 61 GLYSLTSMVTVPSSQKATCNVAHPASSTKVDKTEPIRTPZBPCC-----TCPK 109

QY 187 CPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNK 246
Db 110 CPPENLGGPSVFIFFPKPKDTLMISLTPTVTCVVVDVSDQDEPEVQFTWFVDNKPVGNAE 169

QY 247 TKPREQYNSTYRVVSVLTVHLQDNLNKGKEYCKVSKNKPAPLEKTSKAKGPREQV 306
Db 170 TKPREQYNTTFRVESVLPQHODWLNKGEKCKVKNKLPAPLEKTSKAKGPRMPDV 229

QY 307 YTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--ENNYKTTTPVLSDSGSFFL 364
Db 230 YTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--ENNYKTTTPVLSDSGSFFL 289

QY 365 YSLTVDKSRWQGNVFCSSVMEALHNHYTKSLSPG 404
Db 290 YSLTVDKSAWDQGTVTCTCSVMHEALHNHYTKSAIRSPG 329

Search completed: March 1, 2001, 09:15:49
Job time: 147 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:06 ; Search time 70.93 Seconds
(without alignments)
194.760 Million cell updates/sec

Title: US-09-389-782A-7

Perfect score: 2264

Sequence: 1 ETFFPKYLHYDEETSHQLC.....VMHEALNHNHYTKSLSPG 404

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*

1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1467	64.8	518	15 R51003	Sequence of a reco
2	1460	64.5	485	13 R24016	Fusion protein TNF
3	1332	58.8	376	19 W60037	Antigenic peptide
4	1317	58.2	376	16 W50287	Human Fas antigen
5	1299	57.4	438	16 R81882	Plasmod pbc406/OX4
6	1299	57.4	438	19 W48976	OX40/Fc mutein. C
7	1253	55.3	764	21 Y68949	Fusion protein of
8	1227.5	54.2	375	18 W06683	Heregulin-alpha fu
9	1226	54.2	535	20 Y17414	SVPH1-26 disinteg
10	1225	54.1	963	19 W70540	Integrin beta-1 ch
11	1223.5	54.0	664	19 W71603	Human neurturin re
12	1223.5	54.0	664	21 Y80123	Human NTNR alpha a

13	1221.5	54.0	664	19 W71604	Rat neurturin rece
14	1221.5	54.0	664	21 Y80124	Rat NTNR alpha and
15	1220.5	53.9	400	21 Y15123	Porcine CTLA-4-Ig
16	1217.5	53.8	449	20 W96278	Human noggin/immun
17	1216.5	53.7	388	19 W73513	Rabbit TGFbetaRII:
18	1216.5	53.7	388	21 Y54063	Amino acid sequenc
19	1216	53.7	784	21 Y92207	IL-13/IL-4 dual tr
20	1215.5	53.7	423	21 Y70869	Human interferon-b
21	1215.5	53.7	1158	21 Y92205	Fusion polypeptide
22	1215.5	53.7	1168	21 Y92204	Fusion polypeptide
23	1215	53.7	445	20 Y24153	Bovine LOX-1 extra
24	1214	53.6	592	20 W07097	Human Interleukin-
25	1214	53.6	592	21 Y92185	Human IL-6R-alpha-
26	1212.5	53.6	691	21 Y92202	Fusion polypeptide
27	1212.5	53.6	694	21 Y92203	Fusion polypeptide
28	1212	53.5	388	19 W73514	Human TGFbetaRII:F
29	1212	53.5	388	21 Y54064	Amino acid sequenc
30	1210.5	53.5	387	17 R90920	IL4.Y124D/IgG1 pro
31	1210.5	53.5	408	21 Y44461	Human Interleukin
32	1207.5	53.3	680	15 R48037	tICAM(453)IgG immu
33	1207.5	53.3	859	20 W70796	Human gp130-Fc-His
34	1207.5	53.3	859	21 Y92184	Human gp130-Fc-His
35	1207	53.3	595	20 W86003	Anti-574 single ch
36	1206	53.3	396	18 W18574	Aggrecanase artifi
37	1206	53.3	396	18 W18575	Aggrecanase artifi
38	1205.5	53.2	471	21 Y45030	HUMAN OCRI0-Fc fus
39	1204.5	53.2	482	19 W31646	Human cytokine rec
40	1204	53.2	232	18 W26232	Human IgG1 hinge/F
41	1204	53.2	233	20 Y06617	Human Fc (IgG1).
42	1204	53.2	235	20 Y01372	Amino acid sequenc
43	1204	53.2	254	17 R89441	IgG1 hinge, CH2 an
44	1204	53.2	259	20 Y24154	Protein from pCd51
45	1204	53.2	329	17 R91806	Human immunoglobul

ALIGNMENTS

RESULT 1

R51003
ID R51003 standard; Protein; 518 AA.

XX R51003;

XX 07-OCT-1994 (first entry)

DE Sequence of a recombinant human (rhu) tumour necrosis factor receptor
DE TNFR/fc fusion protein.

XX Tumour necrosis factor receptor; chimeric antibody molecule;
KW immunoglobulin.

XX Synthetic.

XX WO9406476-A.

PD 31-MAR-1994.

PF 14-SEP-1993; 93WO-US08666.

PR 15-SEP-1992; 92US-0946236.

XX (IMMV) IMMUNEX CORP.

XX Jacobs CA, Smith CA;

XX WPI; 1994-118172/14.

XX N-PSDB; Q45225.

PT Treating TNF mediated inflammatory diseases with TNF antagonist -
PT esp. soluble form of TNF receptor, opt. as fusion protein with
PT human immunoglobulin Fc region, esp. for treating arthritis

|||||
 339 tyrvsvltvlhg-wlmgkeyckvsnkalpapiextiskakgpprepvytlppsrdel 397
 317 TKNQVSLTCLVKGFPYSDTAVENSGQPNKYKTPPVLDSDGSPFLYSLKLVVDKSRWQ 376
 398 tknqvsltlcvkgfypsdlavvesngqpennkykttppvldsgsfl-skltvdksrwq 456
 377 QGNVFCVSMHEALHNHYTKLSLSPG 404
 457 qgnvfscvmhealhnhytqkslspsg 484

RESULT 3
 W60037
 ID W60037 standard; Protein; 376 AA.
 AC W60037;
 XX
 XX 11-SEP-1998 (first entry)
 DT
 XX Antigenic peptide hFas (nd29) containing Fc region.
 DE
 XX Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
 KW heart failure; kidney failure; graft-versus-host disease; antibody;
 KW myocardial infarction; ischemic restenosis; endotoxemic shock.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..16
 FT Protein /note= "hFas antigen signal peptide"
 FT Protein 30..376
 FT Protein /note= "hFas (nd29) protein"
 XX
 PN W09818487-A1.
 XX
 XX 07-MAY-1998.
 PD
 XX 31-OCT-1997; 97WO-JP03978.
 PF
 XX 26-SEP-1997; 97JP-0262521.
 PR 31-OCT-1996; 96JP-0290459.
 PR 27-DEC-1996; 96JP-0351718.
 XX
 XX (MOCH) MOCHIDA PHARM CO LTD.
 PA (OSAB-) OSAKA BIOSCIENCE INST.
 XX
 XX Nagata S, Suda T, Yatomi T;
 PI
 XX WPI; 1998-271925/24.
 DR N-PSDB; V34430.
 XX
 XX Use of Fas antagonist for treatment and prevention of
 FT apoptosis-related diseases - such as heart or kidney failure,
 PT graft-versus-host disease or liver disease
 XX
 PS Examples; Fig 5-9; 86pp; Japanese.
 XX
 CC This represents the antigenic peptide hFas (nd29) containing the Fc
 CC region. The invention provides the use of Fas antagonist as an agent for
 CC the treatment and prevention of apoptosis-related diseases. The Fas
 CC antagonist can be a partial Fas antigen peptide containing the
 CC extracellular part of the protein, but lacking the signal sequence, an
 CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is
 CC preferably a humanised antibody. The Fas antagonist is used in the
 CC treatment and prevention of diseases such as myocardial infarction, heart
 CC failure, ischemic heart disease, acute kidney failure, graft-versus-host
 CC disease, ischemic restenosis of the heart, liver or kidney, and
 CC endotoxemic shock, and also as an organ preservative in transplantation.
 CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand
 CC system.
 XX
 XX Sequence 376 AA;
 SQ

Query Match 58.8%; Score 1332; DB 19; Length 376;
 Best Local Similarity 66.7%; Pred. No. 4.3e-77;
 Matches 270; Conservative 15; Mismatches 60; Indels 60; Gaps 11;
 QY 8 LHYDETSQQLCDKPPGTYLKHQCTAKW-KTVCAPCPD-HYYTDSWHTSDCLYCSVP 65
 Db 23 lhhdgqfchk----pcppgerkardctvngdepvcqegkeytdkahfsskrcrr-i 77
 QY 66 CKELQYVKOE--CNRTHNRVCECKEGRYLEIEFCLKHSRCP-----GFGVVGAGTPERNT 119
 Db 78 cdegghgleveinctrtqntkcrckpfnfstvc---ehcdpctkcehglikectltsnt 134
 QY 120 VCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTLTKGNATHDNCISGENSESTOKSGGGG 179
 Db 135 kcke--egrsrsnepks---cdk-----th----- 153
 QY 180 GGTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 239
 Db 154 ---tcpcpapellggsvglfpkpkdtlmisrtpevtcvvvdshedpevkfn 210
 QY 240 VEVHNAKTPREQYNSTYRVSVLVTLVHODWLNKGYCKVSNKALPAPIETISKAG 299
 Db 211 vevhnaaktpreeqynstyrsvsvltvlhqdwlngkeyckvsnkalpapietiskag 270
 QY 300 QPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDTAVENSGQPNKYKTPPVLDSD 359
 Db 271 qprepqvtytlppsrdeltnqvsltlcvkgfypsdlavvesngqpennkykttppvldsd 330
 QY 360 GSPFLYSLKLVTKSRWQGNVFCVSMHEALHNHYTKLSLSPG 404
 Db 331 gsfllyskltvdksrwqgnvfscvmhealhnhytqkslspsg 375
 RESULT 4
 W50287
 ID W50287 standard; Protein; 376 AA.
 XX
 AC W50287;
 XX
 DT 16-JUL-1998 (first entry)
 XX
 DE Human Fas antigen derivative/IgG1 Fc fusion.
 XX
 KW Human; Fas antigen; derivative; apoptosis regulation; gene therapy;
 KW treatment; diabetes; arthritis; lupus; hepatitis; influenza; HIV;
 KW apoptosis modulation; immunoglobulin G1 Fc; IgG1 Fc; fusion.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..16
 FT Peptide /label= sig_peptide
 FT Peptide 17..376
 FT Peptide /label= mat_peptide
 XX
 XX W09742319-A1.
 PN
 XX 13-NOV-1997.
 PD
 XX 01-MAY-1997; 97WO-JP01502.
 PF
 XX 02-MAY-1996; 96JP-0135760.
 PR
 XX (MOCH) MOCHIDA PHARM CO LTD.
 PA (OSAB-) OSAKA BIOSCIENCE INST.
 XX
 XX Nagata S, Nakamura N;
 PI WPI; 1997-558981/51.
 DR N-PSDB; V07004.
 XX

PT Fas antigen derivative containing modified extracellular region -
PT has low antigenicity, promotes apoptosis and is useful in treatment
XX of viral and other diseases
PS Disclosure; Fig 4; 102pp; Japanese.
XX
CC The present sequence is a Fas antigen derivative/IgG1 Fc
CC fusion, which contains a Fas antigen extracellular region lacking
CC one or more amino acid residues in the region from the
CC amino-terminal to (but excluding) the 1st cysteine residue
CC (preferably at least 29 residues are deleted).
CC The derivative is an effective regulator of apoptosis and can be
CC used (either by administration of the polypeptide, or by the use
CC of the coding DNA in gene therapy) to treat a range of diseases,
CC e.g. diabetes, arthritis, lupus and in particular viral diseases
CC such as hepatitis, influenza and HIV, by modulating apoptosis of
CC virus-infected cells.
XX
SQ Sequence 376 AA;

Query Match 58.28; Score 1317; DB 18; Length 376;
Best Local Similarity 65.94; Pred. No. 3.9e-76;
Matches 267; Conservative 16; Mismatches 62; Indels 60; Gaps 11;

QY 8 LHYDETSQLLCDKCPPTGTLKQHTAKW-KTVCAPCPD-HYVTDTSWHTSDECLYCSPV 65
DB 23 lhhdgqfchk----pcppgekrardctvngdepvcpcqeqkeytdkahfskrrcr-l 77

QY 66 CKELQYVQOE--CNRTHNRVCEKRGYLEIEFCLKHSRCPD---GFGVVOAGTPERT 119
DB 78 cdeghgleveinctrtgtkrckpnffcnstvc---ehcdpctkcehgiiiecktiltsnt 134

QY 120 VKCRCPDGFNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCSESTQKSGGGG 179
DB 135 kcke--egrsnepk---cdk-----th----- 153

QY 180 GGGTCCPAPAPELGSPSVFLPPKPKDTLMISRTPEVTVVVDVSHEDPEVKFNWYVDG 239
DB 154 ---tcpcpapellgspsvflppkpkdtlmisrptevtcvvdvshedpevkfnwyvdg 210

QY 240 VEVHNAKTPREEQNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPAPIEKTISKAG 299
DB 211 vevhnaaktpreeqnystyrvsvlvtlqhgdwlngkeyckvsnkalpapiektiskag 270

QY 300 QPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGOPENNYKTTTPVLDSD 359
DB 271 qprepqvylppsrdeltknqvsltcylkgyfspdialvewesngopennykttppvldsd 330

QY 360 GSFFLYSLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPG 404
DB 331 gsfflysltvdksrwqgnvfscsvmhealhnhytkqslspg 375

RESULT 5
R81882
ID R81882 standard; Protein; 438 AA.
XX
AC R81882;
XX
DT 30-MAR-1996 (first entry)
XX
DE Plasmid pBC406/OX40/Fc* encoding an OX40/Fc mutein protein.
XX
KW OX40; OX40-L; cytokine; cell surface molecule; plasmid;
KW pDC406/OX40/Fc*; membrane glycoprotein.
XX
OS Synthetic.
XX
PN US5457035-A.
XX
XX 10-OCT-1995.
XX

PF 23-JUL-1993; 93US-0097827.
XX
PR 23-JUL-1993; 93US-0097827.
XX
PA (IMMV) IMMUNEX CORP.
XX
XX Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
XX PI
XX WPI: 1995-357992/46.
DR N-PSDB; T00829.
XX
XX New isolated DNA encoding the OX40 ligand polypeptide - also vectors
PT and host cells, used to produce recombinant ligand used in e.g.
PT prim. T cell culture, to modulate immune response etc.
XX
XX Example 2; Column 35-38; 26pp; English.
XX
CC This plasmid encodes an OX40/Fc antibody fragment mutein protein,
CC and is used to express a soluble OX40/Fc mutein fusion protein for
CC use in detecting cDNA clones encoding a OX40 ligand. The Fc
CC fragment may be derived from human IgG1, and the plasmid may be
CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell
CC line. Culture supernatant was purified by affinity chromatography
CC and this was used, together with labeled goat anti-human IgG to
CC screen various cell lines.
XX
SQ Sequence 438 AA;

Query Match 57.48; Score 1299; DB 16; Length 438;
Best Local Similarity 60.44; Pred. No. 6.3e-75;
Matches 258; Conservative 22; Mismatches 77; Indels 70; Gaps 6;

QY 20 CDKCPPTGTLKQHTAKWKTVCAPCPDHYTD--SWHTSDECLYCSPCKELQYVQKQECN 77
DB 39 ccrecpqghgmvrchdtrdtclhpcetgyneavnydtckqctqcnh--rsgselkqnc 96

QY 78 RTHNRVCEKRGYLEIEFCLKHSRCPDGFVQVQAGTPERTVCKRCRCPDGFPFNETSKA 137
DB 97 ptqdtvcr-----pgtprqdsqykgldvcpccpgghfs--pgnnq 137

QY 138 PCRKHTNCSVFGLLLTQKGNATHDNCSESTQKSGGGG----- 168
DB 138 ackpwtntctisgkqtrhpasdsldavcedrsllatlletqrptfrptvgtstvwprts 197

QY 169 -----ESTQKSGGGGGGGTCCPCAPAPELLGSPSVFLPPKPKDTLMISRTPEV 217
DB 198 elpstatlveprscdk-----htcpcpapeaeagpsvflppkpkdtlmisrtpev 250

QY 218 TCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQNSTYRVSVLTVLHQDLNKGKEY 277
DB 251 tcvvvdvshedpevkfnwyvdgvevhnaktpreeqnystyrvsvlvtlqhgdwlngkey 310

QY 278 KCKVSNKALPAPIEKTISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAV 337
DB 311 kckvsnkalpapiektiskagqprepqvylppsrdeltknqvsltcylkgyfspdialv 370

QY 338 EWESNGOPENNYKTTTPVLDSDGSFFLYSLTVDKSRWQGNVFCSCVMHEALHNHYTKQ 397
DB 371 ewesngopennykttppvldsdgsfflysltvdksrwqgnvfscsvmhealhnhytkq 430

QY 398 SLSLSPG 404
DB 431 slslspg 437

RESULT 6
W48976
ID W48976 standard; Protein; 438 AA.
XX
AC W48976;
XX
DT 25-SEP-1998 (first entry)
XX

CC within vascular endothelium indicates a role for the polypeptides
 CC in the control of endothelial cell biology. The murine polynucleotide
 CC was identified from a white adipose tissue cDNA library. The polypeptide
 CC is useful for identifying receptors, which bind to and/or are activated
 CC by the polypeptide. The polynucleotide is useful in gene therapy of
 CC cerebral autosomal dominant arteriopathy with subcortical infarcts and
 CC leucoencephalopathy, an autosomal dominant disorder causing ischemic
 CC strokes.

XX Sequence 764 AA;

Query Match 55.3%; Score 1253; DB 21; Length 764;
 Best Local Similarity 58.0%; Pred. No. 9.6e-72;
 Matches 267; Conservative 22; Mismatches 67; Indels 104; Gaps 15;

Qy 11 DEETSHQLLCKCPGTYLKHCHCTAKWTVCA-----PCPDHYY 49
 Db 342 dqensyhcl---cpbg-yvgdqchshlt-cadspcfnggscrerngssyacecpnft 396
 Qy 50 TDSWH-----TSDECLYCSVPCKELQVYKQECNRTHNRVCECKEG-----RYLEIEFCL 98
 Db 397 gsneckkvdrctspcang-----qclnrgpsrtcrpfgtghcelhisdca 446
 Qy 99 KHRSCPPGFGVVGAGT---PERNTVCKRCPDGF-----FSNETSKAPCRKHTNGSV 147
 Db 447 r-spcahg-----gtchdlengpvc-tcpagfgrcevrithidacaspfcngatc-- 496
 Qy 148 FGLLLTQGNATHDNIC-----SGNSESTOKSGGGGGGGTGC 184
 Db 497 -----ytlgpnfnvcncpvgfvgstcefpvgippsfpaapekscdk-----htc 543
 Qy 185 PPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 244
 Db 544 ppcpapellggsfvflppkpkdtlmisrtpevtcvvvdshknpevnfnwvydgvvevnm 603
 Qy 245 AKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREP 304
 Db 604 aktpreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiektiskakgqprep 663
 Qy 305 QVYVLPSPSRDELTKNQVSLTCLVKGFPSDIAVEWESNGQPENNYKTPPVLDSDGSFFL 364
 Db 664 qvytlpsrdeltknqvsitclvkgyfypsdiavewesngqpennykttppvldsdgsffl 723
 Qy 365 YSKITVDKSRWQGNVYFSCVMEALHNHYTQKSLSLSPG 404
 Db 724 yskitvdksrwqgnvfscvmealhnhytcqkslsisp 643

RESULT 8

W06683
 ID W06683 standard; Protein; 375 AA.

XX AC W06683;

XX DT 24-FEB-1997 (first entry)

XX DE Heregulin-alpha fusion protein HRG-alpha-T-Fc.

XX KW Heregulin; epidermal growth factor receptor; EGF; HER4 receptor;
 KW fusion protein; protein phosphorylation; breast cancer;
 KW rHRG-alpha-T-Fc.

XX OS Chimeric Homo sapiens;

XX OS Chimeric synthetic.

XX FH Key Location/Qualifiers

FT Peptide 1..59

FT /label= sig.peptide
 FT /note= "CD5 signal sequence"

FT Misc-difference 26
 FT /note= "x at position 26 corresponds to an
 in-frame stop codon in the coding sequence"

FT Protein 60..375
 FT /label= HRG-alpha-T-Fc
 FT /note= "mature fusion protein produced following
 FT signal peptide cleavage"
 FT Region 60..124
 FT /label= HRG-alpha
 FT /note= "heregulin-alpha EGF-like binding domain"
 FT Cleavage-site 125..141
 FT /label= Thrombin_cleavage_site
 FT Region 142..375
 FT /label= Human Ig-constant region
 XX W09636720-A1.
 XX PN 21-NOV-1996.
 XX PD 14-MAY-1996; 96WO-US06861.
 XX PF 16-MAY-1995; 95US-0441863.
 XX PR (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX PA Aruffo AA, Carlton GW, Culouscou J;
 XX PI WPI; 1997-012095/01.
 XX DR N-PSDB; T45262.
 XX XX Eukaryotic vector encoding growth factor, thrombin cleavage site and
 PT human IgG1 Fc region - esp. encoding EGF-like domain of
 PT heregulin-alpha, -beta3, used for prodn. of recombinant
 PT fusion proteins capable of HER4 receptor activation
 XX Example 1; Fig 7A-B; 188pp; English.
 PS rHRG-alpha-T-Fc (W06683) is a fusion protein that incorporates the
 CC EGF-like binding domain of human heregulin (HRG)-alpha and the Fc
 CC portion of human IgG1. It is the product of a cDNA construct
 CC (T45262) in vector CDW7 and can be produced in eukaryotic (esp.
 CC COS) host cells. The fusion protein, or the HRG-alpha EGF binding
 CC domain produced from it by thrombin cleavage, are both capable of
 CC activating human epidermal growth factor receptor 4 (HER4) and can
 CC be used to study receptor functioning, e.g. protein phosphorylation.
 CC They are also useful in screening assays for (ant)agonists of HER3
 CC and HER4, and bind to cells that express HER4 e.g. MDA-MB-453 human
 CC breast cancer cells, inducing expression of intercellular adhesion
 CC molecule-1.
 XX Sequence 375 AA;

Query Match 54.2%; Score 1227.5; DB 18; Length 375;
 Best Local Similarity 69.9%; Pred. No. 1.8e-70;
 Matches 246; Conservative 7; Mismatches 44; Indels 55; Gaps 7;
 Qy 76 CNRTHNRVCECKEGRYLEIEFCLKHSRCPGFGVVGQAGTPERTVCKRCPDGFSSNETSS 135
 Db 55 clgtsshlvkae---kectfcvnggec---fmvkdlnpsr-ylck-cqpgfigarcte 106
 Qy 136 KAPCRKHTNCSVFGLLLTQGNATHDNICSGNSEST-----OKSGGGGGG----- 180
 Db 107 nvpkmvg-----nqekaelykdp9ggggrlvprfgsgd 142
 Qy 181 -----GGTCPPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 232
 Db 143 pepkscdkthccppcpapefegapsvfllppkpkdtlmisrtpevtcvvvdshdpevk 202
 Qy 233 FNWYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 292
 Db 203 fnwyvdgvevhnaktpreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapi 262
 Qy 293 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFPSDIAVEWESNGQPENNYKTT 352
 Db 263 tiskakgqprepytllppsrdeitknqvsitclvkgyfypsdiavewesngqpennyktt 322

Db 274 ncrasyqvtvs--cpadnyqaclysyagimigdmtpnyvdsptg---ivvspwc----- 323
QY 73 KQECNTRHNRCCKEGRYLEIEF---CLKH-----RSCP--PGFGVVOAGTP 115
Db 324 --scrgsgmmeeee--kfir--dftenpcrlnaiqafngtdvnsvpkpsfqatqprv 378
QY 116 ERNTVCKRCPDGFFSNSTSSKAPCRKHTNCSVFGLLLTQKGNATHD-----NICS 165
Db 379 ekt---pslpddisdstslgtstvtctsvqeggl---kannskelsmofteitltailp 431
QY 166 GNSESTQKSGGGGGGTCTPCPAPPELLGSPSVFLPPKPKDTLMISRTPEVTCVVVDVS 225
Db 432 gprdpvdk-----htcpcpapellggsvfllppkpkdtlmisrtpevtcvvvds 484
QY 226 HEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVVSVLTVLHODWLNGLNGKEYCKVSNKA 285
Db 485 hedpevkfnwyvvgvevhnaktpreeqynstyrvvsvltvlhqdwlngkeyckvsnka 544
QY 286 LPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVRGFYPSDIAVEWESNGQP 345
Db 545 lpapiektiskakgqprepqvyltppsreemtknqvsltclvkgyfypsdiavewesngqp 604
QY 346 ENNYKTTPTPVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSPG 404
Db 605 ennykttptpvlsgsflyskltvdksrwqgnvfscsvmhealhnhtqkslsispg 663

RESULT 12
Y80123
ID Y80123 standard; Protein; 664 AA.
XX
AC Y80123;
XX
DT 19-MAY-2000 (first entry)
XX
DE Human NTNR alpha and IgG fusion protein SEQ ID NO:16.
XX
KW Human; neuturin receptor alpha; NTNR alpha; splenic haematopoiesis;
KW anaemia; thrombocytopaenia; hypoplasia; haemorrhage.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US6025157-A.
XX
PD 15-FEB-2000.
XX
PF 24-OCT-1997; 97US-0957063.
XX
PR 18-FEB-1997; 97US-0038839.
PR 09-JUN-1997; 97US-0049818.
XX
PA (GETH) GENENTECH INC.
XX
PI Hynes MA, Rosenthal A, Klein RD;
XX
DR WPI; 2000-181808/16.
DR N-PSDB; 291459.
XX
PT Isolated nucleic acid molecule encodes a neuturin receptor-alpha amino acid sequence excluding the N-terminal signal peptide -
XX
PS Claim 3; Column 91-96; 78pp; English.
XX
CC The present invention describes a neuturin receptor alpha (NTNR alpha).
CC NTNR alpha binds neuturin. The NTNR alpha nucleic acid molecule is
CC useful for the expression of NTNR alpha, which may be used to identify
CC agonist and antagonist compounds having therapeutic applications, such
CC as enhancing splenic haematopoiesis, treating anaemia,
CC thrombocytopaenia, hypoplasia, or haemorrhage. The present sequence
CC is a NTNR alpha and IgG (immunoglobulin G) fusion protein from the
CC present invention.
XX

SQ Sequence 664 AA;
Query Match 54.0%; Score 1223.5; DB 21; Length 664
Best Local Similarity 62.1%; Pred. No. 6e-70;
Matches 260; Conservative 24; Mismatches 60; Indels 75; Gaps 14;
QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDEGLYCSPVCKELQYV 72
Db 274 ncrasyqvtvs--cpadnyqaclysyagimigdmtpnyvdsptg---ivvspwc----- 323
QY 73 KQECNTRHNRCCKEGRYLEIEF---CLKH-----RSCP--PGFGVVOAGTP 115
Db 324 --scrgsgmmeeee--kfir--dftenpcrlnaiqafngtdvnsvpkpsfqatqprv 378
QY 116 ERNTVCKRCPDGFFSNSTSSKAPCRKHTNCSVFGLLLTQKGNATHD-----NICS 165
Db 379 ekt---pslpddisdstslgtstvtctsvqeggl---kannskelsmofteitltailp 431
QY 166 GNSESTQKSGGGGGGTCTPCPAPPELLGSPSVFLPPKPKDTLMISRTPEVTCVVVDVS 225
Db 432 gprdpvdk-----htcpcpapellggsvfllppkpkdtlmisrtpevtcvvvds 484
QY 226 HEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVVSVLTVLHODWLNGLNGKEYCKVSNKA 285
Db 485 hedpevkfnwyvvgvevhnaktpreeqynstyrvvsvltvlhqdwlngkeyckvsnka 544
QY 286 LPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVRGFYPSDIAVEWESNGQP 345
Db 545 lpapiektiskakgqprepqvyltppsreemtknqvsltclvkgyfypsdiavewesngqp 604
QY 346 ENNYKTTPTPVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSPG 404
Db 605 ennykttptpvlsgsflyskltvdksrwqgnvfscsvmhealhnhtqkslsispg 663

RESULT 13
W71604
ID W71604 standard; Protein; 664 AA.
XX
AC W71604;
XX
DT 19-NOV-1998 (first entry)
XX
DE Rat neuturin receptor alpha/Fc sequence (If2a) fusion protein.
KW Human; neuturin receptor alpha; NTNR-alpha; variant; chimeric;
KW fusion protein; immunoadhesion; ret-expressing cell; neurological;
KW renal; haematological disease.
XX
OS Synthetic.
OS Homo sapiens.
OS Rattus sp.
XX
PN WO9836072-A1.
XX
PD 20-AUG-1998.
XX
PF 17-FEB-1998; 98WO-0503179.
XX
PR 24-OCT-1997; 97US-0957063.
PR 18-FEB-1997; 97US-0802805.
PR 09-JUN-1997; 97US-0871913.
XX
PA (GETH) GENENTECH INC.
XX
PI Hynes MA, Klein RD, Rosenthal A;
XX
XX WPI; 1998-467175/40.
DR N-PSDB; V58007.
PT New polypeptide(s) based on human neuturin receptor alpha and related nucleic acid - useful for increasing survival of


```

AC Y15123;
XX
XX 07-FEB-2000 (first entry)
XX
XX Porcine CTLA-4-Ig construct.
XX
XX Porcine CTLA-4; soluble protein; xenograft; organ transplant; B7; CD28;
KW xenograft-specific immunosuppression; recipient T-cell; anergy;
KW co-stimulatory signal 2; homology; human CTLA-4; bovine CTLA-4.
XX
XX Sus scrofa.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX 162..168
XX Region /label="Flexible_linker
XX /note="Denotes the junction between pCTLA-4"
XX Domain 169..362
XX /label="IgG1_domain
XX Domain
XX
XX WO957266-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-GB01350.
XX
XX 30-APR-1998; 98GB-0009280.
XX
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
XX Lechler IR, Dorling A;
XX
XX WPI; 2000-038815/03.
XX
XX Inhibiting T-cell mediated rejection of xenotransplanted organs
XX
XX Claim 1; Fig 4; 43pp; English.
XX
XX The present sequence is porcine CTLA-4-Ig construct for xenograft
XX -specific immunosuppression. In a pig-to-human transplantation, the
XX soluble protein could comprise the extracellular domain of porcine CTLA-4
XX fused to a human C gamma 1 chain of IgG1. This construct was subcloned
XX into the expression vector pHOOK-3TM and used to transfect DAP.3 or
XX CHO-K1 cells. pCTLA-4-Ig preferentially binds to porcine B7 and blocks
XX its interaction with CD28 on recipient T-cells. This is useful as a
XX species-specific reagent to inhibit human T-cell proliferative responses
XX to a variety of stimulators.
XX
XX Sequence 400 AA;

Query Match 53.9%; Score 1220.5; DB 21; Length 400;
Best Local Similarity 89.5%; Pred. No. 5.2e-70;
Matches 231; Conservative 2; Mismatches 14; Indels 11; Gaps 2;

QY 156 GNATHDNICSGNSESTKSGGGGGG-----TCPCPCAPPELLGGPSVFLPPKPK 206
DB 144 gngtqiyyi--dpepcpdsdggsgaaepkscdkthtcpcpapelligpsvflppkpk 201
QY 207 DTLMSIRTPETCVVVDVSHEDPEVKFNWYDGVGEVHNATKPREOYNSTYRVVSVLT 266
DB 202 dtlmsirtpetcvvvdvshedpevkfnwydgvgevhnahtkpreedynstyrvvsvltv 261
QY 267 LHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCL 326
DB 262 lhqdlngkeyckvsnkalpapiektiskakgprepqvylppsrdeltnqvslctl 321
QY 327 VKGYFPDSIAVWESNGQPNNTTPPVLDSGDSFFLYSKLTVDKSRWQGNVFSQSV 386
DB 322 vkgfypsdiavewesngqpennytppvltdsgdsfflyskltvdksrwqgnvfscsvm 381
QY 387 HEALTHHYTQKSLSPG 404

```

Db 382 healthhytqkslspsg 399

Search completed: March 1, 2001, 09:17:08
Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:20:11 ; Search time 135.68 Seconds
(without alignments)
346.406 Million cell updates/sec

Title: US-09-389-782a-8
Perfect score: 2246
Sequence: 1 MDKTHCTPCPAPPELLGSPS.....OKGNATHDNCISGNSESTQK 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTEMBL_15.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_prodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1009	44.9	401	Q00300	O00300 homo sapien
2	1008	44.9	372	Q9UHP4	Q9UHP4 homo sapien
3	888	39.5	401	O08712	O08712 mus musculus
4	875	39.0	401	O08727	O08727 rattus norv
5	824.5	36.7	437	Q9RIA4	Q9RIA4 mus musculus
6	433	19.3	300	Q95407	Q95407 homo sapien
7	405.5	18.1	302	Q9PUS0	Q9PUS0 salvelinus
8	332.5	14.7	459	O62327	O62327 mus musculus
9	331	14.7	439	Q16042	Q16042 homo sapien
10	327	14.6	482	O08734	O08734 mus musculus
11	287	12.8	655	Q07509	Q07509 homo sapien
12	276	12.3	384	Q9UP60	Q9UP60 homo sapien
13	272.5	12.1	684	Q90544	Q90544 ginglymosto
14	271	12.1	416	Q9NPP6	Q9NPP6 homo sapien
15	252.5	11.2	616	Q9Y6Q6	Q9Y6Q6 homo sapien
16	252.5	11.2	625	O03505	O03505 mus musculus
17	234.5	10.4	348	O057103	O057103 monkeypox v
18	233.5	10.4	349	O057099	O057099 monkeypox v
19	231.5	10.3	349	O057291	O057291 monkeypox v

20	231.5	10.3	349	12	O057100	O057100 monkeypox v
21	231.5	10.3	349	12	O057101	O057101 monkeypox v
22	231.5	10.3	349	12	O057102	O057102 monkeypox v
23	230.5	10.3	348	12	O057277	O057277 monkeypox v
24	230.5	10.3	348	12	O057108	O057108 monkeypox v
25	229	10.2	348	12	O057112	O057112 variola vir
26	229	10.2	348	12	O05407	O05407 variola vir
27	226.5	10.1	349	12	O057284	O057284 camelopox vi
28	226.5	10.1	349	12	O057098	O057098 camelopox vi
29	226	10.1	349	12	O057110	O057110 variola vir
30	226	10.1	349	12	O057111	O057111 variola vir
31	226	10.1	349	12	O089118	O089118 variola vir
32	226	10.1	349	12	O089098	O089098 variola vir
33	223	9.9	350	12	O057116	O057116 cowpox viru
34	222.5	9.9	349	12	O057097	O057097 camelopox vi
35	222.5	9.9	355	12	O083308	O083308 cowpox viru
36	221	9.8	349	12	O057109	O057109 variola vir
37	220.5	9.8	349	12	O057305	O057305 cowpox viru
38	217.5	9.7	326	12	O057120	O057120 cowpox viru
39	217.5	9.7	326	12	O057122	O057122 cowpox viru
40	217.5	9.7	351	12	O073559	O073559 cowpox viru
41	217.5	9.7	360	12	O057118	O057118 cowpox viru
42	216.5	9.6	351	12	O057117	O057117 cowpox viru
43	214.5	9.6	350	12	O057123	O057123 cowpox viru
44	212.5	9.5	347	12	O057115	O057115 cowpox viru
45	211.5	9.4	351	12	O057121	O057121 cowpox viru

ALIGNMENTS

RESULT	1
O00300	
ID	O00300 PRELIMINARY; PRT; 401 AA.
AC	O00300; O60236;
DT	01-JUL-1997 (TREMBLrel. 04, Created)
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE	OSTEOPTERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE	(OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER IIB).
GN	TNFRSF11B OR OPG OR OCIF.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=KIDNEY;
RX	MEDLINE=97262071; PubMed=9108485;
RA	Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA	Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA	Derose M., Elliott R., Colombero A., Tan H.-L., Trall G., Sullivan J.,
RA	Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA	Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA	Suggs S., Boyle W.J.;
RT	"Osteoprotegerin: a novel secreted protein involved in the regulation
RT	of bone density.";
RL	Cell 89:309-319(1997).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=LUNG FIBROBLAST;
RX	MEDLINE=98151033; PubMed=9492069;
RA	Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA	Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
RA	Tsuda E., Morinaga T., Higashio K.;
RT	"Identity of osteoclastogenesis inhibitory factor (OCIF) and
RT	osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RT	osteoclastogenesis in vitro.";
RL	Endocrinology 139:1329-1337(1998).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=PLACENTA;
RX	MEDLINE=98351569; PubMed=9688283;

RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human
RL Eur. J. Biochem. 254:685-691(1998).
CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; AB002146; BAA25910.1; -.
DR EMBL; AB008822; BAA32076.1; -.
DR EMBL; AB008821; BAA32076.1; JOINED.
DR EMBL; U94332; AAB53709.1; -.
DR HSP; P25942; ICDF.
DR MIM; 602643; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 183 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 44.9%; Score 1009; DB 4; Length 401;
Best Local Similarity 99.4%; Pred. No. 3.7e-80;
Matches 173; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 228 KETPPKYLHYDEETSHQLLCKDPCPGTYLKQHCTAKWTKVACPCPDHYTDSWHTSDEC 287
Db :|||||
QY 288 LYCSPVKELQYVQECNTRNHRVCECKEGRYLEIEFCLKHSRCPGPGVVOAGTTPERT 347
Db 81 LYCSPVKELQYVQECNTRNHRVCECKEGRYLEIEFCLKHSRCPGPGVVOAGTTPERT 140
QY 348 VCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQK 401
Db 141 VCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQK 194

RESULT 2
Q9UHP4 PRELIMINARY; PRT; 372 AA.
ID -Q9UHP4

AC Q9UHP4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).
DR EMBL; AF134187; AAF20168.1; -.
DR HSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
FT NON_TER 1
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 44.9%; Score 1008; DB 4; Length 372;
Best Local Similarity 100.0%; Pred. No. 4.2e-80;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 ETFFPKYLHYDEETSHQLLCKDPCPGTYLKQHCTAKWTKVACPCPDHYTDSWHTSDECL 288
Db 1 ETFFPKYLHYDEETSHQLLCKDPCPGTYLKQHCTAKWTKVACPCPDHYTDSWHTSDECL 60
QY 289 YCSPVKELQYVQECNTRNHRVCECKEGRYLEIEFCLKHSRCPGPGVVOAGTTPERTV 348
Db 61 YCSPVKELQYVQECNTRNHRVCECKEGRYLEIEFCLKHSRCPGPGVVOAGTTPERTV 120
QY 349 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQK 401
Db 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQK 173

RESULT 3
O08712 PRELIMINARY; PRT; 401 AA.
ID O08712; 070202;
AC O08712; 070202;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF).
GN TNFRSF11B OR OPG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=KIDNEY;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.Q., Woodson A., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
of bone density";
RL Cell 89:309-319(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA, AND NIH SWISS;
RX MEDLINE=98382527; PubMed=9714833;

RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,
 RA Higashio K.;
 RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
 RL gene and its expression in embryogenesis.";
 RL Gene 215:339-343(1998).
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
 CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
 CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
 CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
 CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,
 CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND
 CC PLACENTA. NOT DETECTED IN SPLEEN.
 CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT
 CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY
 CC 15 TO DAY 17.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL; U94331; AAB53708.1; -.
 DR EMBL; A9013898; BAA28269.1; -.
 DR EMBL; A9013903; BAA33388.1; -.
 DR EMBL; A9013899; BAA33388.1; JOINED.
 DR EMBL; A9013900; BAA33388.1; JOINED.
 DR EMBL; A9013901; BAA33388.1; JOINED.
 DR EMBL; A9013902; BAA33388.1; JOINED.
 DR HSSP; P25942; ICDF.
 DR MGD; MGI:109587; Op9.
 DR INTERPRO; IPR000488; -.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 3.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -. 1.
 KW Glycoprotein; Repeat; Cytokine; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 401 OSTEOPROTEGERIN.
 FT DOMAIN 23 201 4 X TNFR-CYS.
 FT REPEAT 23 63 TNFR-CYS 1.
 FT REPEAT 64 106 TNFR-CYS 2.
 FT REPEAT 107 143 TNFR-CYS 3.
 FT REPEAT 144 201 TNFR-CYS 4.
 FT DOMAIN 306 365 DEATH DOMAIN.
 FT DISULFID 41 54 BY SIMILARITY.
 FT DISULFID 44 62 BY SIMILARITY.
 FT DISULFID 65 80 BY SIMILARITY.
 FT DISULFID 83 97 BY SIMILARITY.
 FT DISULFID 87 105 BY SIMILARITY.
 FT DISULFID 118 142 BY SIMILARITY.
 FT DISULFID 145 160 BY SIMILARITY.
 FT CARBOHYD 98 98 POTENTIAL.
 FT CARBOHYD 165 165 POTENTIAL.
 FT CARBOHYD 178 178 POTENTIAL.
 FT CARBOHYD 289 289 POTENTIAL.
 FT VARIANT 138 138
 FT VARIANT 161 161
 FT VARIANT 165 165
 FT VARIANT 165 165
 FT VARIANT 288 288
 FT VARIANT 296 296
 FT SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;
 SQ
 Query Match 39.5%; Score 888; DB 11; Length 401;
 Best Local Similarity 86.8%; Pred. No. 1.4e-69;
 Matches 151; Conservative 8; Mismatches 15; Indels 0; Gaps 0;
 QY 228 KETFPKYLHYDETSQQLCDKCPGTYLKQHCATKWKTCVPCPDHYTDSWHTSDEC 287

Db 21 QETLPPKYLHYDETSQQLCDKCAPGTYLKQHCATKWKTCVPCPDHYTDSWHTSDEC 80
 QY 288 LYCSPVCKELQYKQECNRTHNRVCECKEGRYLEIEFCLKHSRCPGFGVVGAGTPRNT 347
 Db 81 VYCSPVCKELQYKQECNRTHNRVCECKEGRYLEIEFCLKHSRCPGFGVVGAGTPRNT 140
 QY 348 VCKRCPDGFSSNETSSKAPCRKHTNCVFGLLLTOKGNATHDNCISGNSSTOK 401
 Db 141 VCKKCPDGFSSNETSSKAPCRKHTNCVFGLLLTOKGNATHDNCISGNSSTOK 194
 RESULT 4
 O08727
 ID O08727 PRELIMINARY; PRT; 401 AA.
 AC O08727;
 DT 01-JUL-1997 (TEMBLrel. 04, Created)
 DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
 DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
 DE (OCIF)
 GN TNFRSF11B OR OPG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=INTESTINE;
 RX MEDLINE=97262071; PubMed=9108485;
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
 RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
 RA Derose M., Elliott R., Colombero A., Tan H.-L., Trall G., Sullivan J.,
 RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
 RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
 RA Suggett S., Boyle W.J.;
 RT "Osteoprotegerin: a novel secreted protein involved in the regulation
 RT of bone density.";
 RL Cell 89:309-319(1997).
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
 CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
 CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
 CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
 CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL; U94330; AAB53707.1; -.
 DR HSSP; P25942; ICDF.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -. 1.
 KW Glycoprotein; Repeat; Cytokine; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 401 OSTEOPROTEGERIN.
 FT DOMAIN 23 201 4 X TNFR-CYS.
 FT REPEAT 23 63 TNFR-CYS 1.
 FT REPEAT 64 106 TNFR-CYS 2.
 FT REPEAT 107 143 TNFR-CYS 3.
 FT REPEAT 144 201 TNFR-CYS 4.
 FT DOMAIN 306 365 DEATH DOMAIN.
 FT DISULFID 41 54 BY SIMILARITY.
 FT DISULFID 44 62 BY SIMILARITY.
 FT DISULFID 65 80 BY SIMILARITY.
 FT DISULFID 83 97 BY SIMILARITY.
 FT DISULFID 87 105 BY SIMILARITY.
 FT DISULFID 118 142 BY SIMILARITY.
 FT DISULFID 145 160 BY SIMILARITY.
 FT CARBOHYD 98 98 POTENTIAL.
 FT CARBOHYD 165 165 POTENTIAL.

Q95407	PRELIMINARY;	PRT;	300 AA.
AC	Q95407		
ID	Q95407		
DT	01-MAY-1999 (TrEMBLrel. 10, Created)		
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)		
DE	DECAY RECEPTOR 3 (M68) (M68C) (M68E).		
GN	DCR3 OR TR6.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=99087326; PubMed=9872321;		
RX	Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,		
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,		
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,		
RA	Goddard A.D., Botstein D., Ashkenazi A.;		
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and		
RT	colon cancer.";		
RL	Nature 396:699-703(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE=BLOOD;		
RC	MEDLINE=99253915; PubMed=10318773;		
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;		
RT	"A newly identified member of tumor necrosis factor receptor		
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";		
RN	J. Biol. Chem. 274:13733-13736(1999).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE=PANCREAS;		
RC	MEDLINE=20122600; PubMed=10655513;		
RX	Bai C., Connolly B., Metzger M.L., Hilliard C.A., Liu X., Sandig V.,		
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;		
RT	"Overexpression of M68/DCR3 in human gastrointestinal tract tumors		
RT	independent of gene amplification and its location in a four-gene		
RT	cluster.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).		
RL	EMBL; AF104419; RAD03056.1; -		
DR	EMBL; AF134240; RAD29688.1; -		
DR	EMBL; AF217796; AAF35244.1; -		
DR	EMBL; AF217793; AAF33685.1; -		
DR	EMBL; AF217794; AAF33686.1; -		
DR	HSP; P25942; ICDF.		
DR	INTERPRO: IPR000361; -		
DR	INTERPRO: IPR001368; -		
DR	PFAM: PF00020; TNFR_C6; 4.		
DR	PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.		
DR	PROSITE: PS01186; EGF_2; UNKNOWN_1.		
DR	PROSITE: PS50050; TNFR_NGFR_2; 2.		
DR	PRODOM: PD000771; -; 1.		
DR	Receptor.		
KW	SEQUENCE 300 AA; 32679 MW; F90AE33718449AF CRC64;		
QY	Query Match	19.3%;	Score 433; DB 4; Length 300;
QY	Best Local Similarity	42.9%;	Pred. No. 4.9e-30;
Db	Matches 69; Conservative 29; Mismatches 63; Indels 0; Gaps		
QY	233 PKYLHDEETSHQLDKCPGPGTYLKQHCATKWKTVACAPDHYHTDSMHTSDECLYCSP	292	
Db	34 PTPYWRDAETGERLVCAQCPCPGTFVQRPCCRDSPTTCGCPPRHYTQFWNYLCRCYCNV	93	
QY	293 VKELQYVQECNFRNHRVCEKEGRYLEIEFCLKHRSCPPGFGVVGQAGTPERNVTKRC	352	
Db	94 LGEREEREAACHATHNRACRRTGFFAHAGFCLHSCPPGAGVIATGPTSQNTQCP	153	
QY	353 PGGFFNETSSKAPCRKHNCSVFGLLLTQKNATHDNICS	393	
Db	154 PPGTFSSSSSQEQCPHNCNTALGIALNVPSSSHDTLCT	194	

Best Local Similarity 31.4%; Pred. No. 3.3e-16;
Matches 80; Conservative 33; Mismatches 110; Indels 32; Gaps 10;

QY 3 KTHTCP-----PCPAPELLGGPSVLFPPKPK-----DTLMISRTPEVT 41
Db 119 KHTNPSQDVTPCPVPSTPTPTSPST-PTTPSPSCCHPRLSLHRPALEDLLILGSEANLT 177

QY 42 CVVVDVSHEDPEVKFNMYDGVVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGKEYK 101
Db 178 CTLTGL-RDASGVTFWTTPSSGR--SAVOGPPERDLGGCVSVSSVLPGCAEPNHNKTKTF 234

QY 102 KVSNNKALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQ-VSLTCLVKGYFSPSDIAV 160
Db 235 CTAAYPESTPLTATLSKS-GNTRFPEVHLLPPSEELALNELVTLTCLARGSPKDVLY 293

QY 161 EWESNGQ--PENNYKTTTPVLD-SDG--SFFLYSKLTVDKSRWQGNVFCSCVYHNAH 215
Db 294 RWLQSGQELPREKYLFWASRQSPGTTTFAVTISILRVAEDWKKGDTFSCWVGHEALPL 353

QY 216 HYTKSLSPGKRET 230
Db 354 AFTQKTIIDRLAGKPT 368

RESULT 13
QY0544 PRELIMINARY; PRT; 684 AA.

AC QY0544;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RX MEDLINE=95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks.";
RL Nature 374:168-173(1995).
DR EMBL: U18701; AAB48195.1; -.
DR HSSP: P01857; 1FC1.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 1g; 6.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_3.
KW Signal.
FT SIGNAL. 1 18 POTENTIAL.
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 12.1%; Score 272.5; DB 13; Length 684;
Best Local Similarity 33.3%; Pred. No. 1.3e-15;
Matches 71; Conservative 35; Mismatches 98; Indels 9; Gaps 7;

QY 20 SVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPE-VKFNMYDGVVHNAKTPREEQYN 78
Db 459 SVSLLRP-PEETWTQOTATVEIV---YSDLENIKVFWQVGVKRRKGVETQNPENWG 514

QY 79 STYRVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAK-GOPREQVYTLPPSRD 137
Db 515 SKSTIVSKLVAMASEWDSGTEYVCLVEDSELPTPVKASIRKANVSQMHPPKVLLHPSTD 574

QY 138 EL-TKNQVSLTCLVKGYFSPSDIAVESNGQ-PENNYKTTTPVLDSDGSPFLYSKLTVDK 195
Db 575 EIDTENSATLMCLATNFHPAIEIVGVWMANDTLSDSGYRTQVDSEKSGSGSFVTDRLRLTA 634

QY 196 SRWQGNVFCSCVYHNAHNYTKSLSLSPGK 228
Db 635 AEWNSTDTTSCVGVHPSL-NRDLIRSTNKSNGK 666

RESULT 14
QYNNPP6 PRELIMINARY; PRT; 416 AA.

ID QYNNPP6;
AC QYNNPP6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansoorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RT "The European IMAGE consortium for integrated molecular analysis of
RT human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: ALJ389978; CAB97534.1; -.
FT NON_TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 12.1%; Score 271; DB 4; Length 416;
Best Local Similarity 32.8%; Pred. No. 1e-15;
Matches 77; Conservative 34; Mismatches 112; Indels 12; Gaps 9;

QY 2 DKHTGCPCPAPELLGGPSVLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 61
Db 172 DVTVPVPPPPCC-HPRLSLHRPALED-LLIGSEANLTCTLGL-RDASGATFTWTPS 228

QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAK 421
Db 229 SGK--SAVOGPPERDLGGCVSVSSVLPGCAQPNWHGFTCTAAHPKELPTLANITKS- 285

QY 122 GQPREQVYTLPPSRDELTKNQ-VSLTCLVKGYFSPSDIAVESNGQ--PENNYKTTTPV 178
Db 286 GNTRFPEVHLLPPSEELALNELVTLTCLARGSPKDVLRWLQSGQELPREKYLTVASR 345

QY 179 LD-SDG--SFFLYSKLTVDKSRWQGNVFCSCVYHNAHNYTKSLSLSPGKRET 230
Db 346 QEPFSGTTFVAVTISILRVAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPT 400

RESULT 15
QY606 PRELIMINARY; PRT; 616 AA.

ID QY606;
AC QY606;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE RECEPTOR ACTIVATOR OF NF-KAPPA-B PRECURSOR (TNF-RELATED ACTIVATION-
DE INDUCED CYTOKINE RECEPTOR) (RANK).
GN TNFRSF1A OR RANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98032977; PubMed=93671155;
RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
RA Tometsko M.E., Roux E.R., Teepe M.C., DuBoise R.F., Cosman D.,

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:48 ; Search time 40.97 Seconds
(without alignments)
316.083 Million cell updates/sec

Title: US-09-389-782A-8
Perfect score: 2246
Sequence: 1 MDKTHTCPPCAPPELLGPGS.....CKGNATHDNCISGNSSTQK 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	Length	DB ID	Description
1	1233	54.9	330	1	GCL_HUMAN	P01857 homo sapien
2	1142.5	50.9	326	1	GC2_HUMAN	P01859 homo sapien
3	1135	50.5	327	1	GC4_HUMAN	P01861 homo sapien
4	1126	50.1	320	1	GC3_HUMAN	P01860 homo sapien
5	918.5	40.9	323	1	GC_RABIT	P01870 oryctolagus
6	899	39.6	329	1	GC2_CAVPO	P01862 cavia porce
7	845.5	37.6	329	1	GC3_MOUSE	P22436 mus musculus
8	838	37.3	333	1	GC3_RAT	P20761 rattus norv
9	835.5	37.2	398	1	GC3_MOUSE	P03987 mus musculus
10	818.5	36.4	326	1	GCL_RAT	P20759 rattus norv
11	817.5	36.4	324	1	GCL_MOUSE	P01868 mus musculus
12	812.5	36.2	393	1	GC1_MOUSE	P01869 mus musculus
13	809.5	36.0	329	1	GCC_RAT	P20762 rattus norv
14	809	36.0	330	1	GCAA_MOUSE	P01863 mus musculus
15	804	35.8	399	1	GCAM_MOUSE	P01865 mus musculus
16	802	35.7	335	1	GCAB_MOUSE	P01864 mus musculus
17	785.5	35.0	322	1	GCA_RAT	P20760 rattus norv
18	779	34.7	336	1	GCB_MOUSE	P01866 mus musculus
19	774	34.5	405	1	GCEB_MOUSE	P01867 mus musculus
20	362	16.1	454	1	MUC_HUMAN	P01871 homo sapien
21	361	16.1	455	1	MUC_MOUSE	P01872 mus musculus
22	360	16.0	391	1	MUCB_HUMAN	P04220 homo sapien
23	359	16.0	421	1	EPC_MOUSE	P06336 mus musculus
24	358	15.9	429	1	EPC_RAT	P01855 rattus norv
25	352	15.7	428	1	EPC_HUMAN	P01854 homo sapien
26	347	15.4	476	1	MUCM_MOUSE	P01873 mus musculus
27	346	15.4	458	1	MUC_RABIT	P03988 oryctolagus
28	336	15.0	450	1	MUC_CANFA	P01874 canis famil
29	336	15.0	454	1	MUC_MESAU	P06337 mesocricetu
30	333	14.8	457	1	MUC_SUNMU	P20768 sunus muri
31	333	14.8	461	1	TNR2_HUMAN	P20333 homo sapien
32	332.5	14.8	474	1	TNR2_MOUSE	P25119 mus musculus
33	332	14.8	479	1	MUCM_RABIT	P04221 oryctolagus

RESULT 1

ID	GCL_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG GAMMA-1 CHAIN C REGION.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
RN	[3]			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
RN	[4]			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The			
RT	peptides and peptides of the H-chain, alignment of the tryptic			
RT	peptides and peptides of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X. Intrachain disulfide bonds."			

P23085 heterodontu
P01879 oryctolagus
P23087 heterodontu
P01875 gallus gall
P27312 mus musculu
P23086 heterodontu
P23088 heterodontu
P23084 heterodontu
P01876 homo sapien
P20758 gorilla gor
P01877 homo sapien
P25942 homo sapien

34 303 13.5 438 1 HVC2_HETFR
35 299.5 13.3 299 1 ALC_RABIT
36 293.5 13.1 438 1 HVC3_HETFR
37 287.5 12.8 446 1 MUC_CHICK
38 284 12.6 289 1 CD40_MOUSE
39 282.5 12.6 393 1 HVC3_HETFR
40 278.5 12.4 461 1 HVC4_HETFR
41 278 12.4 370 1 HVC1_HETFR
42 276 12.3 353 1 ALC1_HUMAN
43 271.5 12.1 353 1 ALC1_GORGO
44 271 12.1 340 1 ALC2_HUMAN
45 264 11.8 277 1 CD40_HUMAN

ALIGNMENTS

Biochemistry 9:3188-3196(1970).
 [7]
 RN DISULFIDE BONDS.
 RX MEDLINE=77070267; PubMed=1002129;
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT cyanogen bromide cleavage products, and the disulfide bridges.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE=81208100; PubMed=7236608;
 RA Deisenhofer J.;
 RT "Crystallographic refinement and atomic models of a human Fc fragment
 RT and its complex with fragment B of protein A from Staphylococcus
 RT aureus at 2.9- and 2.8-A resolution.";
 RL Biochemistry 20:2361-2370(1981).
 CC -!- MISCELLANEOUS: NIE HAS THE G1M(17) ALLOTYPIC MARKER, 97-K, & THE
 CC G1M(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE G1M(3)
 CC MARKER & THE G1M (NON-1) MARKERS.
 CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
 CC 35,116,198,269 & 272.
 CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
 CC 268-272.
 CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
 CC RESIDUES 198,267&272.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J00228; AAC82527.1; ALT_INIT.
 DR PIR; A02146; GHU.
 DR PDB; 1FC1; 15-JUL-92.
 DR PDB; 1FC2; 15-JUL-92.
 DR MIN; 147100; -.
 DR INTERPRO; IPR000495; -.
 DR INTERPRO; IPR003006; -.
 DR PFAM; PF00047; Ig; 3.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 3D-structure.
 FT NON_TER 1 1
 FT DOMAIN 1 98 CH1.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 223 CH2.
 FT DOMAIN 224 330 CH3.
 FT DISULFID 27 83
 FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).
 FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).
 FT DISULFID 144 204
 FT DISULFID 250 308
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .).
 FT VARIANT 97 97 K -> R (IN G1M(3) MARKER).
 FT VARIANT 239 239 D -> E (IN G1M(NON-1) MARKER).
 FT VARIANT 241 241 /FTId=VAR_003886.
 FT VARIANT 330 330 /FTId=VAR_003887.
 FT STRAND 123 126 L -> M (IN G1M(NON-1) MARKER).
 FT STRAND 130 134 /FTId=VAR_003888.
 FT TURN 136 137 REMOVED POST-TRANSLATIONALLY.
 FT STRAND 141 148
 FT STRAND 158 162
 FT TURN 163 164

FT STRAND 165 166
 FT STRAND 175 178
 FT STRAND 183 190
 FT HELIX 193 197
 FT TURN 198 199
 FT STRAND 202 206
 FT STRAND 215 219
 FT STRAND 227 227
 FT STRAND 230 234.
 FT HELIX 238 240
 FT TURN 241 242
 FT STRAND 245 256
 FT STRAND 260 266
 FT TURN 267 268
 FT STRAND 269 270
 FT STRAND 274 276
 FT STRAND 280 281
 FT TURN 283 284
 FT STRAND 287 296
 FT STRAND 297 301
 FT HELIX 302 303
 FT TURN 306 312
 FT TURN 313 314
 FT TURN 316 317
 FT STRAND 320 324
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;
 Query Match 54.9%; Score 1233; DB 1; Length 330;
 Best Local Similarity 100.0%; Pred. No. 1.2e-78;
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 DKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 61
 DB 104 DKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 163
 QY 62 GVEVHNARKTPREQYNSTYRVSVLTIVLHODWLNKREYKCKVSNKALPAPIETISKAK 121
 DB 164 GVEVHNARKTPREQYNSTYRVSVLTIVLHODWLNKREYKCKVSNKALPAPIETISKAK 223
 QY 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
 DB 224 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 283
 QY 182 DGSFELYSLKLVDRKRWQOGNVSFSCSVNHEALHNHYTQKSLSLSPGK 228
 DB 284 DGSFELYSLKLVDRKRWQOGNVSFSCSVNHEALHNHYTQKSLSLSPGK 330
 RESULT 2
 GC2_HUMAN
 ID GC2_HUMAN STANDARD; PRT; 326 AA.
 AC P01859;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-2 CHAIN C REGION.
 GN IGHG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82197621; PubMed=6804948;
 RA Ellison J.W., Hood L.E.;
 RT "Linkage and sequence homology of two human immunoglobulin gamma
 RT heavy chain constant region genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
 RN [2]
 RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
 RX MEDLINE=81007873; PubMed=6774012;
 RA Wang A.-C., Tung E., Fudenberg H.H.;
 RT "The primary structure of a human IgG2 heavy chain: genetic,

evolutionary, and functional implications.";
J. Immunol. 125:1048-1054(1980).
[3]
SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
MEDLINE=80001357; PubMed=113060;
Connell G.E., Parr D.M., Hofmann T.;
"The amino acid sequences of the three heavy chain constant region
domains of a human IgG2 myeloma protein.";
Can. J. Biochem. 57:758-767(1979).
[4]
SEQUENCE OF 238-275 (ZIE).
MEDLINE=80114419; PubMed=118920;
Hofmann T., Parr D.M.;
"A note of the amino acid sequence of residues 381-391 of human
immunoglobulins gamma chains.";
Mol. Immunol. 16:923-925(1979).
[5]
REVIEWS TO 25; 59; 60 AND 264-268 (ZIE).
Hofmann T., Parr D.M.;
Submitted (MAR-1980) to the PIR data bank.
[6]
SEQUENCE OF 1-121 (DOT).
MEDLINE=95355298; PubMed=7737190;
Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
"Characterization of the two unique human anti-flavin monoclonal
immunoglobulins.";
Eur. J. Biochem. 228:886-893(1995).
[7]
DISULFIDE BONDS.
MEDLINE=72033500; PubMed=4940472;
Milstein C., Frangione B.;
"Disulphide bridges of the heavy chain of human immunoglobulin G2.";
Biochem. J. 121:217-225(1971).
[8]
DISULFIDE BONDS.
MEDLINE=69064124; PubMed=5782707;
Frangione B., Milstein C., Pink J.R.L.;
"Structure studies of immunoglobulin G.";
Nature 221:145-148(1969).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; V00554; CAA23814.1; -
EMBL; V00554; CAA23815.1; -
EMBL; V00554; CAA23816.1; -
EMBL; V00554; CAA23817.1; -
PIR; A02148; G2HU.
MIM; 147110; -
INTERPRO; IPR000495; -
INTERPRO; IPR003006; -
PFAM; PF00047; ig; 3.
PROSITE; PS00290; IG_MHC; 2.
Immunoglobulin domain; Immunoglobulin C region.
NON_TER 1 1
DOMAIN 1 98 CHI.
DOMAIN 99 110 HINGE.
DOMAIN 111 219 CH2.
DOMAIN 220 326 CH3.
INTERCHAIN (WITH A LIGHT CHAIN).
DISULFID 14 14
DISULFID 27 83
DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 140 200
DISULFID 246 304
VARIANT 60
S -> A (IN MYELOMA PROTEINS TIL & ZIE).

FT SITE 156 156 /FTid=VAR_003889.
FT MOD_RES 326 326 AT OR NEAR THE COMPLEMENT-BINDING SITE.
FT SEQUENCE 326 AA; 35884 MW; 8310876C6878CF9C CRC64;
SQ
Query Match 50.9%; Score 1142.5; DB 1; Length 326;
Best Local Similarity 94.1%; Pred. No. 2.1e-72;
Matches 209; Conservative 8; Mismatches 4; Indels 1; Gaps 1;
QY 7 CPCPCAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 56
Db 106 CPCPCAPP-VAGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVH 64
QY 67 NAKTKPREEOYNSTYRVVSVLTVLDHDLNGKEYCKVSNKALPAPIEKTISKAKGPRE 26
Db 165 NAKTKPREEQFNSTFRVSVLTVVHDLNGKEYCKVSNKGLPAPIEKTISKAKGPRE 24
QY 127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPVLDSDGSF 186
Db 225 PQVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPVLDSDGSF 284
QY 187 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTOKLSLSPGK 228
Db 285 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTOKLSLSPGK 326
RESULT 3
GC4_HUMAN STANDARD; PRT; 327 AA.
ID AC P01861;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-4 CHAIN C REGION.
GN IGHG4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83157104; PubMed=6299662;
RA Ellison J.W., Buxbaum J.N., Hood L.E.;
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL DNA 1:11-18(1981).
RN [2]
RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE=70207560; PubMed=4192699;
RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; K01316; AAB59394.1; ALT_INIT.
PIR; A02150; G4HU.
MIM; 147130; -
INTERPRO; IPR000495; -
INTERPRO; IPR003006; -
PFAM; PF00047; ig; 3.
PROSITE; PS00290; IG_MHC; 2.
Immunoglobulin domain; Immunoglobulin C region.
NON_TER 1 1
DOMAIN 1 98 CHI.
DOMAIN 99 110 HINGE.
DOMAIN 111 220 CH2.
DOMAIN 111 220

Best Local Similarity 90.3%; Pred. No. 2.6e-71;
Matches 205; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 2 DKTHTCPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 64 DTPPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVQFKWYVD 123

QY 62 GVEVHNKATKPREQNSYRVSVTLVHODWLNKGEYKCKVSNKALPAPIETISKAK 121
Db 124 GVQVHNKATKPREQNSYRVSVTLVHODWLNKGEYKCKVSNKALPAPIETISKYK 183

QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181
Db 184 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 243

QY 182 DGSFFLYSKLTVDKSWQOGNFSCVMHEALHNHYTQKSLSLSPGK 228
Db 244 DGSFFLYSKLTVDKSWQOGNFSCVMHEALHNHYTQKSLSLSPGK 290

RESULT 5
GC_RABIT STANDARD; PRT; 323 AA.
AC F01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA CHAIN C REGION.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit igg heavy chain from the recombinant
RT F-1 haplotype."
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RT immunoglobulin G of different allotype."
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit igg: isolation of a cDNA encoding gamma
RT heavy chain and identification of two genomic C gamma genes."
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461105;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RT immunoglobulin G."
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16426; AAA31289.1; -.
DR PIR; A02161; GHRB.
DR INTERPRO; IPR000495; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> E (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).
FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 698AA118D579A8B CRC64;

Query Match 40.9%; Score 918.5; DB 1; Length 323;
Best Local Similarity 71.7%; Pred. No. 6.7e-57;
Matches 167; Conservative 29; Mismatches 32; Indels 5; Gaps 2;

QY 1 MDKT---HTC--PPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 55
Db 91 VDKTAPSTCTSKPTCPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 150

QY 56 FNNYDGVVHNKATKPREQNSYRVSVTLVHODWLNKGEYKCKVSNKALPAPIEK 115
Db 151 FTWYINNEQVTRAPPLREQQFNSTIRVSTLPTITHODWLNKGEYKCKVSNKALPAPIEK 210

QY 116 TISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 175
Db 211 TISKAGQPLEPKVYTMGPPRELSRSVSLTCTMNGFYFSDISVEWENKGAEDNKTT 270

QY 176 PPVLDSGDSFFLYSKLTVDKSWQOGNFSCVMHEALHNHYTQKSLSLSPGK 228
Db 271 PAVLDSGDSFFLYSKLTVDKSWQOGNFSCVMHEALHNHYTQKSLSRSPGK 323

RESULT 6
GC2_CAVPO STANDARD; PRT; 329 AA.
ID GC2_CAVPO
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2 CHAIN C REGION.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Hystricognathi; Caviidae; Cavia.
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshstein B.K., Hussain Q.Z., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the


```

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR INTERPRO: IPR000495; -
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 37.3%; Score 838; DB 1; Length 333;
Best Local Similarity 66.4%; Pred. No. 2.6e-51;
Matches 148; Conservative 33; Mismatches 42; Indels 0; Gaps 0;

QY 6 TCPCPAPPELLGSPVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 65
DB 111 TCHKCPPELLGSPVFIFFPKPKDILLISQNAKVTCTVVVDVSEEPDVGQFSWVNVEV 170
QY 66 HNAKTKPREOYNSTYRVVSVLTVLDWLNKGYCKYKSNKALPAPIEKTISKAKGQPR 125
DB 171 H7AQTQPREQYNSTYRVVSVLTVLDWLNKGYCKYKSNKALPAPIEKTISKAKGLVR 230
QY 126 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSF 185
DB 231 KPQVYVNGPTEQLTEQTVSLTCLTSGFLPNDIGVEMTSNGHTEKYNKTEPVWDSGDSF 290
QY 186 FLYSKLTVDKSRWQGNVFCSCVNMHEALHNHYTQKSLSLSPGK 228
DB 291 FMYSKLVNRSWRDRAFPVCSVHVEGLHHHVEKSISSPPGK 333

RESULT 9
CC3M_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=85027161; PubMed=6092053;
RA Wells J.A.; Word C.J.; Rimm D.; Der-Balan G.P.; Martinez H.M.;
RA Tucker P.W.; Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
[2]
RN SEQUENCE OF 328-398 FROM N.A.
RP MEDLINE=84041483; PubMed=6314258;
RA Komaromy M.; Clayton L.; Rogers J.; Robertson S.; Kettman J.;
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
-----
DR EMBL: J00451; AAB59655.1; -
DR EMBL: V01526; CAA24767.1; ALT_SEQ.
DR PIR: A02155; G3MSM.
DR INTERPRO: IPR000495; -
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT DOMAIN 345 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 37.2%; Score 835.5; DB 1; Length 398;
Best Local Similarity 66.2%; Pred. No. 4.7e-51;
Matches 149; Conservative 34; Mismatches 39; Indels 3; Gaps 1;

QY 8 PP---CPAPPELLGSPVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 64
DB 106 PGSSCPPGNGILGSPVFIFFPKPKDLMISLTPKVTCTVVVDVSEDDPDVHVSFVDNKE 165
QY 65 VNAKTKPREOYNSTYRVVSVLTVLDWLNKGYCKYKSNKALPAPIEKTISKAKGQPR 124
DB 166 VHTAQTQPREQYNSTYRVVSVLTVLDWLNKGYCKYKSNKALPAPIEKTISKAKGQPR 225
QY 125 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 184
DB 226 QTPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 285
QY 185 FFLYSLKLTVDKSRWQGNVFCSCVNMHEALHNHYTQKSLSLSPGKE 229
DB 286 YFLYSLKLTVDKSRWQGNVFCSCVNMHEALHNHYTQKSLSLSPGKE 330

RESULT 10
GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR: PS0017; PS0017.
DR INTERPRO: IPR000495; -
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
```

```
FT NON_TER 1 1 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 36.4%; Score 818.5; DB 1; Length 326;
Best Local Similarity 63.1%; Pred. No. 5.6e-50;
Matches 142; Conservative 41; Mismatches 35; Indels 7; Gaps 2;

QY 7 CPGCPAPPELLGG---PSVLEPPPKPKDLMISRTPEVTCVVVDVSHEDPEVKENWYVDGV 63
DB 106 CKPC-----ICGSEVSSVFIFPPKPKDVLTLTPKVKTCVVVDISQDDPEVHFSWFDVV 161
QY 64 EVHNAKTPREEQYNSTYRVSVLTFLHQLDNLGKEYCKVSNKALPAPIETISKAKGQ 123
DB 162 EVHTAQTAPPEEQFNSTFRSVSELPILHQLDNLGTRFCKVTSAAFPSPIEKTIKSPGGR 221
QY 124 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLDSDG 183
DB 222 TVPHVYMTSPTEKTEQNEVSIICMVKGFYPPDIYEWQMGQPNQENYKNTPTPTMDTG 281
QY 184 SFFLYSKLTVDKSRWQOGNFCFSVMHEALHNHYTQKSLSLSPGK 228
DB 282 SYFLYSLNVLNKKERKQOGNTFTCSVLHLEGLNHHNTEKLSLSPGK 326

RESULT 11
GCL_MOUSE
ID GCL_MOUSE STANDARD; PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
SEQUENCE FROM N.A.
MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RT Cell 18:559-568(1979).
[2]
SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RA MEDLINE=80202559; PubMed=6769752;
RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
[3]
SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RA MEDLINE=80012837; PubMed=113776;
RA Rogers J., Clarke P., Salsner W.;
RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
heavy chain.";
RL Nucleic Acids Res. 6:3305-3321(1979).
[4]
SEQUENCE (MYELOMA PROTEIN MOPC 21).
```

```
RX MEDLINE=78242288; PubMed=98524;
RA Adetugbo K.;
RT "Evolution of immunoglobulin subclasses. Primary structure of a
murine myeloma gammal chain.";
RL J. Biol. Chem. 253:6068-6075(1978).
[5]
RN DISULFIDE BONDS (MOPC 21).
RP MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00795; CAA24176.1; -
DR PIR; A02159; GIMS.
DR MGD; MGI:96446; Igh-4.
DR INTERPRO; IPR000495; -
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1 97
FT DOMAIN 98 110 CH1.
FT DOMAIN 111 217 HINGE.
FT DOMAIN 218 324 CH2.
FT DISULFID 27 82 CH3.
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 36.4%; Score 817.5; DB 1; Length 324;
Best Local Similarity 62.5%; Pred. No. 6.6e-50;
Matches 140; Conservative 44; Mismatches 35; Indels 5; Gaps 2;

QY 7 CPP--CPAPELLGGSVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKENWYVDGVE 64
DB 104 CKPCICVPEV---SSVFIFPPKPKDVLTLTPKVKTCVVVDISQDDPEVQFSWFVDVDE 160
QY 65 VHNAKTPREEQYNSTYRVSVLTFLHQLDNLGKEYCKVSNKALPAPIETISKAKGQP 224
DB 161 VHTAQTAPPEEQFNSTFRSVSELPILHQLDNLGTRFCKVTSAAFPAPIETIKTKGRP 220
QY 125 REPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLDSDGS 184
DB 221 KAPQVYTIPTPPKQMAKDKVSLTCMTIDTFEEDITVEMQWNGQPAENYKNTQPTMTNGS 280
QY 185 FFYLSKLTVDKSRWQOGNFCFSVMHEALHNHYTQKSLSLSPGK 228
DB 281 YFVYSLNVLNKKERKQOGNTFTCSVLHLEGLNHHNTEKLSLSPGK 324
```

RESULT 12
GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.;
RL Cell 18:559-568(1979).
[2]
RN SEQUENCE OF 323-393 FROM N.A.
RP MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowan A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain".
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
[3]
RN SEQUENCE OF 323-366 FROM N.A.
RP MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains".
RL Cell 26:19-27(1981).
[4]
RN SEQUENCE OF 1-44 FROM N.A.
RP MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains".
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
SEGMENT OF MU CHAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; V00793; CAA241172.1; -
CC EMBL; V00793; CAA241173.1; -
CC EMBL; V00793; CAA241174.1; -
CC PIR; B02159; GIMSM.
CC MGD; MGI:96446; IGH-4.
CC INTERPRO: IPR000495; -
CC INTERPRO: IPR003006; -
CC PFAM; PF00047; ig; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1 1
FT DOMAIN 98 110 CH1.
FT DOMAIN 111 217 HINGE.
FT DOMAIN 218 324 CH2.
FT DOMAIN 218 324 CH3.

FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 138 N-LINKED (GLCNAC. . .).
FT CARBOHYD 174 174 POTENTIAL.
FT DISULFID 244 302 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 340 357
FT DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7AICE27 CRC64;

Query Match 36.2%; Score 812.5; DB 1; Length 393;
Best Local Similarity 62.3%; Pred. No. 1.8e-49;
Matches 139; Conservative 44; Mismatches 35; Indels 5; Gaps 2;

QY 7 CPP--CPAPELLGGPSVFLFPKPKDITLMTISRTPEVTCVVVYDVSHEDPEYKFNMYVDGVE 64
Db 104 CKPCICTVPEV---SSVFIPPKPKDVLITLTPKVTCTVVVDISKDDPEVQFSWEVDVE 160

QY 65 VHNAKTKPREQVNSTYRVSVTLVHLDWLNKGYCKVSKNKPALPAPIETISKAKQP 124
Db 161 VHTAQTQPREQFNSTFRSVSELPIMHQDWLNKGYCKVSKNKPALPAPIETISKAKGRP 220

QY 125 REPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEVNESGQENNYKTPPVLDSDGS 184
Db 221 KAFQVYTIPTPKQMAKDKVSLTCTMTDFEDITVEWQNGQPAENYKNTQPIINTNGS 280

QY 185 FFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPG 227
Db 281 YFYVSKLVNQSWEAGNTFTCSVLHGLHNHHTKSLSHSPG 323

RESULT 13
GCC,RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE IG GAMMA-2C CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
region cDNA: extensive homology to mouse gamma 3".
RL Eur. J. Immunol. 18:317-319(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X07189; CAA30169.1; -
CC PIR; S00847; S00847.
CC INTERPRO: IPR000495; -
CC INTERPRO: IPR003006; -
CC PFAM; PF00047; ig; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 98 97 CH1.
FT DOMAIN 113 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.

FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 82
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 143 203
 FT DISULFID 249 307
 SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 36.0%; Score 809.5; DB 1; Length 329;
 Best Local Similarity 63.8%; Pred. No. 2.4e-49;
 Matches 143; Conservative 41; Mismatches 37; Indels 3; Gaps 1;

Qy 8 PP---CPAPELGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVFNWVDGVE 64
 Db 106 PPTDICSCDDNLGRPSVFIFPPKPKDILMTLPKVTCTVVDVSEEDVQFSEFVNDVR 165

Qy 65 VHNATKPREQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIETISKAKGP 124
 Db 166 VFTAQTOPHEQLNGTFRVWSTLHQHDMWSGKEFKCKVNNKDLPSPIETISKPRGKA 225

Qy 125 REPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVESNGQPNKYKTPPVLDSDGS 184
 Db 226 RTPQYITIPPREQMSKNKVSITCMVTGFYPASISVERNGELEDQDYKNTLPVLDSDS 285

Qy 185 FFLYKSLTVDRSKRWQGVNFCVSMHEALHNHYTKSLSPGK 228
 Db 286 YFLYKSLSDVDTSMRGDIYTCVSVVHEALHNHYTKSLSPGK 329

RESULT 14
 GCAA_MOUSE STANDARD; PRT; 330 AA.
 AC P01863;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE IG GAMMA-2A CHAIN C REGION, A ALLELE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8107654; PubMed=6777755;
 RA Sikorav J.-L., Auffray C., Rougeon F.;
 RT "Structure of the constant and 3' untranslated regions of the murine
 RT Balb/c gamma 2a heavy chain messenger RNA.";
 RL Nucleic Acids Res. 8:3143-3155(1980).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81198976; PubMed=6262729;
 RA Yanawaki-Kataoka Y., Miyata T., Honjo T.;
 RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
 RT and evolution of heavy chain genes: further evidence for intervening
 RT sequence-mediated domain transfer.";
 RL Nucleic Acids Res. 9:1365-1381(1981).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81223894; PubMed=6787604;
 RA Ollo R., Auffray C., Morchamps C., Rougeon F.;
 RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
 RT suggests that exons can be exchanged between genes in a multigenic
 RT family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
 RN [4]
 RP MYELOMA PROTEIN MOPC 173.
 RX MEDLINE=74175517; PubMed=4831970;
 RA Bourgois A., Fougereau M., Rocca-Serra J.;
 RT "Determination of the primary structure of a mouse IgG2a
 RT immunoglobulin: amino-acid sequence of the FC fragment. Implications
 RT for the evolution of immunoglobulin structure and function.";
 RL Eur. J. Biochem. 43:423-435(1974).
 RN [5]

RP DISULFIDE BONDS.
 RA MEDLINE=73056887; PubMed=4565406;
 RX de Preval C., Fougereau M.;
 RT "Determination of the primary structure of a mouse gamma G2a
 RT immunoglobulin. Identification of the disulfide bridges.";
 RL Eur. J. Biochem. 30:452-462(1972).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V00798; CAA24178.1; -;
 DR PIR; A02152; G2MSA.
 DR INTERPRO; IPR000495; -;
 DR INTERPRO; IPR003006; -;
 DR PFAM; PF00047; Ig; 3.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 82
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 144 204
 FT DISULFID 250 308
 FT MOD_RES 330 330
 SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 36.0%; Score 809; DB 1; Length 330;
 Best Local Similarity 65.2%; Pred. No. 2.6e-49;
 Matches 146; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

Qy 7 CPP---CPAPELGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVFNWVDGVE 64
 Db 107 CPPCKPAPNLGGPSVFIFPPKPKDILMTLPKVTCTVVDVSEDDPDVQLSWFNWVE 166

Qy 65 VHNATKPREQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIETISKAKGP 124
 Db 167 VHTAQTOPHEQYNSTLRVSVLTVLHODWLNKGYCKVSNKALPAPIETISKAKGP 226

Qy 125 REPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVESNGQPNKYKTPPVLDSDGS 184
 Db 227 RAPQVYVLPPEEEMTKKQVTLTCMTDMPEDIYVETNNGKTELKNTPEVLDSDGS 286

Qy 185 FFLYKSLTVDRSKRWQGVNFCVSMHEALHNHYTKSLSPGK 228
 Db 287 YFLYKSLTVDRSKRWQGVNFCVSMHEALHNHYTKSLSPGK 330

RESULT 15
 GCAA_MOUSE STANDARD; PRT; 399 AA.
 AC P01865;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG GAMMA-2A CHAIN C REGION, MEMBRANE-BOUND FORM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82222190; PubMed=6283537;
 RA Yanawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
 RT "Nucleotide sequences of gene segments encoding membrane domains of
 RT immunoglobulin gamma chains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).

```
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END. ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00471; AAB59661.1; ALT_INIT.
DR PIR; A02154; G2MSAM.
DR MGI; 96443; IGH-1.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR003006; -.
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;
```

Query Match 35.8%; Score 804; DB 1; Length 399;
Best Local Similarity 65.0%; Pred. No. 7.2e-49;
Matches 145; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

```
QY 7 CPP--CPAPELGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 CPPCKCPAPNLLGGPSVFIKPKIKDVLMSLSPIVTCVVVDVSEDDPDVQISWVNNVE 166
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 VHNAKTKPREEQNSTYRVYSVLTVLHQLNGKGYCKVKYNKALPAPIEKTISKAKGQP 124
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 167 VHTAQQTREDYNSLTVVVSALPIQHDWMSGKFKACKNNKDLPAPIERTISKPGSV 226
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 125 REPQVYTLPPSRDLTKNQVSLTCLVKGYFSPDAVWESNGQPENNYKTPPVLDSDGS 184
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 227 RAFPQVYVLPPEEEMTKQVTLTCMTDFWPDIIYVEWTNNGKTELNYKNTPEVLDSDGS 286
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 185 FFLYSKLTVDKSRWQGNVFCSCVMHEALNHHYTKQSLSPG 227
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 287 YFMYSKLVEKKNVERNYSYCSVHGLEHNNHHTKFSRTPG 329
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: March 1, 2001, 09:17:49
Job time: 267 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:49 ; Search time 79.26 Seconds
(without alignments)
343.530 Million cell updates/sec

Title: US-09-389-782A-8
Perfect score: 2246
Sequence: 1 MDKTHTCPPCPAPPELLGGPS.....QKNATHDNCISGNSESTQK 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 6790655 residues
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	54.9	255	4 S31866	Ig gamma-1 chain C
2	1233	54.9	330	1 GHUU	Ig gamma-1 chain C
3	1227	54.6	374	2 S69339	Ig heavy chain V r
4	1227	54.6	374	2 S72664	Ig heavy chain V r
5	1180	52.5	234	2 PT0207	Ig gamma chain C r
6	1146	51.0	377	2 A23511	Ig gamma-3 chain C
7	1144	50.9	377	2 A60764	Ig gamma-3 chain C
8	1142.5	50.9	326	1 G2HU	Ig gamma-2 chain C
9	1135	50.5	327	1 G4HU	Ig gamma-4 chain C
10	1121	49.9	289	1 G3HUI	Ig gamma-3 heavy C
11	918.5	40.9	323	1 GHRB	Ig gamma chain C r
12	906.5	40.4	328	2 I47160	Ig gamma 2b chain
13	906.5	40.4	328	2 I47159	Ig gamma 2a chain
14	903.5	40.2	277	2 I47162	Ig gamma 4 chain C
15	889	39.6	329	1 G2GP	Ig gamma-2 chain C
16	885.5	39.4	328	2 I47158	Ig gamma 1 chain c
17	878.5	39.1	328	2 I47161	Ig gamma 3 chain c
18	855.5	38.1	470	2 S22080	Ig heavy chain pre
19	846	37.7	308	2 C30554	Ig heavy chain C r
20	846	37.7	472	2 S31459	Ig gamma-1 chain -
21	845.5	37.6	329	1 G3MSC	Ig gamma-3 chain C
22	838	37.3	333	2 PS0018	Ig gamma-2b chain
23	835.5	37.2	398	1 G3MSM	Ig gamma-3 chain C
24	827.5	36.8	444	2 PC4436	monoclonal antibody
25	818.5	36.4	326	2 PS0017	Ig gamma-1 chain C
26	817.5	36.4	324	1 G1MS	Ig gamma-1 chain C
27	812.5	36.2	393	1 G1MSM	Ig gamma-1 chain C
28	809.5	36.0	329	2 S00847	Ig gamma-2c chain
29	809	36.0	330	1 G2MSA	Ig gamma-2a chain

30 809 36.0 469 2 S37483
31 804 35.8 399 1 G2MSAM
32 802 35.7 335 1 G2MSAB
33 794 35.4 446 2 S40295
34 785.5 35.0 322 2 PS0019
35 779 34.7 336 1 G2MS11
36 779 34.7 474 2 S25057
37 774 34.5 405 1 G2MSBM
38 764 34.0 327 2 S06611
39 757 33.7 475 2 S01321
40 707 31.5 180 2 I46732
41 577.5 25.7 249 2 S69340
42 574.5 25.6 218 2 A36040
43 571 25.4 152 2 S14236
44 395.5 17.6 572 2 B46529
45 362 16.1 343 2 S25644

ALIGNMENTS

RESULT 1
S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C:Accession: S31866

R:Filpula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene produc

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FIL>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: Immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 54.9%; Score 1233; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 9.6e-72;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKHTCCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVYVDSHEDPEVFNWYVD 61

Db 29 DKHTCCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVYVDSHEDPEVFNWYVD 88

QY 62 GVEVHNKATKPREQYNTSVRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETISKAK 121

Db 89 GVEVHNKATKPREQYNTSVRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETISKAK 148

QY 122 GQPREQVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181

Db 149 GQPREQVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 208

QY 182 DGSFPLYSKLTVDKSRQOGNMFSCSVNHEALHNHYTQKSLSLSPGK 228

Db 209 DGSFPLYSKLTVDKSRQOGNMFSCSVNHEALHNHYTQKSLSLSPGK 255

RESULT 2
GHUU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999

C:Accession: A93433; S36863; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene

A:Reference number: A93433; MUID:82274238

A:Accession: A93433

A:Molecule type: DNA
A:Residues: 1-330 <ELL>
A:Cross-references: EMBL:Z117370
A:Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,
A:Note: Lys-330 is removed after translation
R:Harris, L.J.
submitted to the EMBL Data Library, October 1992
A:Reference number: S33904
A:Accession: S36861
A:Molecule type: DNA
A:Residues: 2-330 <HAR>
A:Cross-references: EMBL:Z117370
R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of a
A:Reference number: S33887; MUID:83001943
A:Accession: S33887
A:Molecule type: DNA
A:Residues: 88-113:235-330 <TAK>
A:Cross-references: EMBL:Z117370
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,
Biochemistry 9, 3161-3170, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequence
A:Reference number: A90563; MUID:71064024
A:Contents: myeloma protein Eu
A:Accession: B90563
A:Molecule type: protein
A:Residues: 1-96,'R',98-135 <CUN>
A:Note: this sequence has the G1m(3) marker, 97-Arg
R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequence
A:Reference number: A90564; MUID:71064025
A:Contents: Eu
A:Accession: A90564
A:Molecule type: protein
A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',240,
A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met
R:Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nle).
igen Primärstruktur.
A:Reference number: A91668; MUID:77070269
A:Contents: myeloma protein Nie
A:Accession: B91668
A:Molecule type: protein
A:Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27
A:Note: this sequence has the G1m(17) and G1m(1) markers
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL
A:Reference number: A91723; MUID:83289131
A:Contents: myeloma protein KOL; disulfide bonds
A:Accession: A91723
A:Molecule type: protein
A:Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH
A:Note: this sequence has the G1m(3) and G1m(non-1) markers
R:Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfide
A:Reference number: A90565; MUID:71064027
A:Contents: annotation; disulfide bonds
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A:Reference number: A91667; MUID:77070267
A:Contents: annotation; disulfide bonds
C:Genetics
A:Gene: GDB:IGHG1
A:Cross-references: GDB:120085; OMIM:147100
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 114/1; 224/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/Disulfide bonds: #status experimental
F:103/Disulfide bonds: interchain (to light chain) #status experimental
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 54.9%; Score 1233; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHCPCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 104 DKTHCPCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 163
QY 62 GVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 121
Db 164 GVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 223
QY 122 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTTPVLD 181
Db 224 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTTPVLD 283
QY 182 DGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPCK 228
Db 284 DGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPCK 330

RESULT 3
S69339
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-2000
C:Accession: S69339
R:Khamlich, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A:Reference number: S69339; MUID:95262687
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 54.6%; Score 1227; DB 2; Length 374;
Best Local Similarity 99.1%; Pred. No. 3.5e-71;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHCPCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 148 DKTHCPCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 207
QY 62 GVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 121
Db 208 GVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 267
QY 122 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTTPVLD 181
Db 268 GQPREPQVYTLPPSRDEMTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTTPVLD 327
QY 182 DGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPCK 228
Db 328 DGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKSLSPCK 374

QY 182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 221

Db 151 DTPPCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYD 210
Qy 62 GVEVHNATKKPREOYNSTYRVVSVLTVTLHODWLNKGVCKVSNKALPAPIETKTSKAK 121
Db 211 GVEVHNATKKPREOYNSTYRVVSVLTVTLHODWLNKGVCKVSNKALPAPIETKTSKTK 270
Qy 122 GQPREPQVYVTPPSPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181
Db 271 GQPREPQVYVTPPSPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 330
Qy 182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSPGK 228
Db 331 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSPGK 377

RESULT 8
G2HU
Ig gamma-2 chain C region - human
C:Species: Homo sapiens (man)
C:Date: 30-Apr-1981 #sequence_revision 13-Jun-1983 #text_change 21-Jul-2000
C:Accession: A93906; A92809; A90752; A93132; A02148
R:Ellison, J.; Hood, L.
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982
A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con
A:Reference number: A93906; MUID:82197621
A:Accession: A93906
A:Molecule type: DNA
A:Residues: 1-326 <ELL>
A:Cross-references: GB:V00554; GB:J00230; NID:932759; PIDN:CAB58438.1; PID:g6066056
A:Note: Lys-326 is probably removed posttranslationally
R:Wang, A.C.; Tung, E.; Fudenberg, H.H.
J. Immunol. 125, 1048-1054, 1980
A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f
A:Reference number: A92809; MUID:81007873
A:Contents: myeloma protein T11
A:Accession: A92809
A:Molecule type: protein
A:Residues: 1-19, 'Q', '21-57', 'Z', '59', 'A', '61-193', 'D', '195-325 <WAN>
A:Note: Trp-156 is at or near the complement-binding site
R:Connell, G.E.; Parr, D.M.; Hofmann, T.
Can. J. Biochem. 57, 758-767, 1979
A:Title: The amino acid sequences of the three heavy chain constant region domains of a
A:Reference number: A90752; MUID:80001357
A:Contents: myeloma protein Zie
A:Accession: A90752
A:Molecule type: protein
A:Residues: 1-24, 'E', '26-57', 'EV', '60-85', '132-171', '222', '175', 'B', '177-193', 'D', '195-196', 'Q', '198-
A:Note: This sequence has since been revised
R:Hofmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g
A:Reference number: A93132; MUID:80114419
A:Contents: Zie
A:Accession: A93132
A:Molecule type: protein
A:Residues: 238-275 <HOF>
R:Hofmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A:Reference number: A94591
A:Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268
A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati
ned
R:Milstein, C.; Frangione, B.
Biochem. J. 121, 217-225, 1971
A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.
A:Reference number: A90253; MUID:72033500
A:Contents: annotation; myeloma protein Sa, disulfide bonds
R:Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A:Title: Structural studies of immunoglobulin G.
A:Reference number: A93157; MUID:69064124
A:Contents: annotation; Sa, disulfide bonds

C:Genetics:
A:Gene: GDB:IGHG2
A:Cross-references: GDB:119338; OMIM:147110
A:Map position: 14q32.33-14q32.33
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:133-202/Domain: immunoglobulin homology <IM2>
F:239-306/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:27-83,140-200,246-304/Disulfide bonds: #status experimental
F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.9%; Score 1142.5; DB 1; Length 326;
Best Local Similarity 94.1%; Pred. No. 7.2e-66;
Matches 209; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

Qy 7 CPCPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYDGVH 66
Db 106 CPCPCAPP-VAGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYDGVH 164
Qy 67 NAKTKPREEQNSTYRVVSVLTVTLHODWLNKGVCKVSNKALPAPIETKTSKAGQPRE 126
Db 165 NAKTKPREEQNSTYRVVSVLTVTLHODWLNKGVCKVSNKALPAPIETKTSKAGQPRE 224
Qy 137 PQVYVTPPSPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFF 186
Db 225 PQVYVTPPSPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFF 284
Qy 187 LYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSPGK 228
Db 285 LYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSPGK 326

RESULT 9
G4HU
Ig gamma-4 chain C region - human
C:Species: Homo sapiens (man)
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999
C:Accession: A90933; A90249; A02150
R:Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A:Reference number: A90933; MUID:83157104
A:Accession: A90933
A:Molecule type: DNA
A:Residues: 1-327 <ELL>
A:Note: the sequence was determined from the germline gene
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constan
A:Reference number: A90249; MUID:70207560
A:Accession: A90249
A:Molecule type: protein
A:Residues: 1-30;81-326 <PIN>
C:Genetics:
A:Gene: GDB:IGHG4
A:Cross-references: GDB:119340; OMIM:147130
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 111/1; 221/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:99-110/Region: hinge
F:134-203/Domain: immunoglobulin homology <IM2>
F:240-307/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:27-93,141-201,247-305/Disulfide bonds: #status predicted
F:106,109/Disulfide bonds: interchange (to heavy chain) #status experimental
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.5%; Score 1135; DB 1; Length 327;
Best Local Similarity 93.7%; Pred. No. 2.2e-65;
Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 7 CPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
DB 106 CPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 165

QY 67 NAKTKPREQYNSTYRVSVLTVLDHQLNGKEYCKVSNKALPAPIETKTSKAKQPRE 136
DB 166 NAKTKPREQYNSTYRVSVLTVLDHQLNGKEYCKVSNKALPAPIETKTSKAKQPRE 225

QY 127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSF 186
DB 226 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSF 285

QY 187 LYSKLVTDKSRWQGNVSCSVSMHEALHNHYTOKLSLSLSPGK 228
DB 286 LYSKLVTDKSRWQGNVSCSVSMHEALHNHYTOKLSLSLSPGK 327

RESULT 10
G3HUWI
Ig gamma-3 heavy chain disease proteins - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C:Accession: A90442; A92219; A90198; A93915; A02149
R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy chain
A:Reference number: A90442; MUID:81021548
A:Contents: heavy chain disease protein Wis
A:Accession: A90442
A:Molecule type: protein
A:Residues: 1-289 <PRA>
A:Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchange
A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
A:Note: the sequence of residues 42-76 was taken from the reference that follows
R:Michaelson, T.E.; Frangione, E.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
A:Reference number: A92219; MUID:77118561
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W
A:Accession: A92219
A:Molecule type: protein
A:Residues: 12-97 <MIC>
A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter
R:Wolfeinstein-Tedel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
A:Reference number: A90198; MUID:77021516
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues
A:Accession: A90198
A:Molecule type: protein
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>
A:Note: this protein lacks most of the V region, all of the CH1 region, and part of the
R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
A:Reference number: A93915; MUID:82247835
A:Contents: heavy chain disease protein Omn
A:Accession: A93915
A:Molecule type: mRNA
A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157
A:Note: a carboxyl-terminal Lys is removed posttranslationally
A:Note: this sequence may represent an allelic form or another gamma chain subclass

C:Comment: The heavy chain disease protein Wis is shown.

C:Genetics:
A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
F:203-270/Domain: immunoglobulin homology <IMM>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 49.9%; Score 1121; DB 1; Length 289;
Best Local Similarity 90.3%; Pred. No. 1.5e-64;
Matches 204; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 2 DKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
DB 64 DKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 123

QY 62 GVEVHNKTKPREQYNSTYRVSVLTVLDHQLNGKEYCKVSNKALPAPIETKTSKAK 121
DB 124 GVEVHNKTKPREQYNSTYRVSVLTVLDHQLNGKEYCKVSNKALPAPIETKTSKAK 183

QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181
DB 184 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 243

QY 182 DGSFFLYSKLTVDKSRWQGNVSCSVSMHEALHNHYTOKLSLSLSPG 227
DB 244 DGSFFLYSKLTVDKSRWQGNVSCSVSMHEALHNHYTOKLSLSLSPG 289

RESULT 11
GHRB
Ig gamma chain C region - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 6-Jul-1999
C:Accession: A91749; A90290; A93928; A90245; A94416; A02161
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I hap
A:Reference number: A91749; MUID:84030930
A:Accession: A91749
A:Molecule type: mRNA
A:Residues: 1-323 <BER>
A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-
Biochem. J. 151, 337-349, 1975
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunog
A:Reference number: A90290; MUID:76135459
A:Accession: A90290
A:Molecule type: protein
A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>
R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
A:Title: Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma heavy ch
A:Reference number: A93928; MUID:83299917
A:Accession: A93928
A:Molecule type: mRNA
A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>
A:Cross-references: GB:M16426; NID:gl65111; PIDN:AAA31289, 1; PID:gl65112
A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic ma
R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
Biochem. J. 116, 249-259, 1970
A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul
A:Reference number: A90245; MUID:70110015
A:Accession: A90245
A:Molecule type: protein
A:Residues: 132-143, 'E', 145-161 <FRU>
R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
In Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wikse
A:Reference number: A94416


```
F:248-308/Disulfide bonds: #status experimental

Query Match
Best Local Similarity 79.6%; Score 889; DB 1; Length 329;
Matches 162; Conservative 24; Mismatches 36; Indels 2; Gaps 1;

QY 6 TCPCPAPELLGGPSVFLFPPKPKDTLMIISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 65
Db 106 TCPKCPPENLGGPSVFIFPPKPKDTLMIISLTTPVTCVVVDVSDPEQVFTFVFNKPV 165
QY 66 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPR 125
Db 166 GNAETKPREQYNTFFVESYLP IQHODWLGRGKFKCVYNKALPAPIEKTIISKAKGR 225
QY 126 EPQVYTLPPSRDELTNQVSLTCLVKGFPYSDIAVESNGQP--ENNYKTTTPVLDSDG 183
Db 226 MPDVYTLPPSRDELSKSRVSTCLIIINFFADIHVESNRVPSRSEKYNTPPIEDADG 285
QY 184 SFFLYSKLITDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPG 227
Db 286 SYFLYSKLITVDKSAWDQGVTVVTCSSVMHEALHNHYTQKAISRSPG 329

Search completed: March 1, 2001, 09:15:50
Job time: 148 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:08 ; Search time 70.93 Seconds
(without alignments)
193.313 Million cell updates/sec

Title: US-09-389-782A-8
Perfect score: 2246
Sequence: 1 MDKTHCTPCPAPPELLGPPS.....QKGNATHDNCNGNSESTQK 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36.*
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1239.5	55.2	652	19 W48650	Heavy chain of hma
2	1238	55.1	374	19 W83963	Recombinant human
3	1238	55.1	374	19 W49075	Recombinant human
4	1237	55.1	235	20 Y01372	Amino acid sequenc
5	1236.5	55.1	651	18 W26649	Chimeric receptor
6	1236.5	55.1	692	18 W26650	Chimeric receptor
7	1234	54.9	347	13 R27163	CD2 binding LFA-3-
8	1234	54.9	347	21 Y83136	Human transmembran
9	1234	54.9	388	19 W73513	Rabbit TGFbetaRII:
10	1234	54.9	388	19 W73514	Human TGFbetaRII:
11	1234	54.9	388	21 Y54063	Amino acid sequenc
12	1234	54.9	388	21 Y54064	Amino acid sequenc

13	1234	54.9	399	21 Y70867	Human interferon-b
14	1234	54.9	418	21 Y70868	Human interferon-b
15	1234	54.9	423	21 Y70869	Human interferon-b
16	1234	54.9	446	15 R58753	VCAM 2D-IgG. Homo
17	1234	54.9	446	20 Y23986	VCAM 2D-IgG, a sol
18	1234	54.9	446	20 Y01037	VCAM 2D-IgG protei
19	1234	54.9	446	20 W96743	A VCAM 2D-IgG1 fus
20	1234	54.9	482	19 W31646	Human cytokine rec
21	1233	54.9	232	18 W26232	Human IgG1 hinge/F
22	1233	54.9	233	20 Y06617	Human Fc (IgG1).
23	1233	54.9	259	20 Y24154	Protein from pCD51
24	1233	54.9	329	17 R91806	Human immunoglobul
25	1233	54.9	351	14 R43685	Human kappa immuno
26	1233	54.9	371	10 P91918	Sequence of the li
27	1233	54.9	371	10 P93558	Linkered human IgG
28	1233	54.9	376	19 W60037	Antigenic peptide
29	1233	54.9	379	19 W83962	Recombinant human
30	1233	54.9	379	19 W49073	Recombinant human
31	1233	54.9	379	19 W49074	Recombinant human
32	1233	54.9	387	17 R90920	IL4 Y124D/IgG1 pro
33	1233	54.9	396	18 W18574	Aggrecanase artifi
34	1233	54.9	396	18 W18575	Aggrecanase artifi
35	1233	54.9	400	21 Y15123	Porcine CTLA-4-Ig
36	1233	54.9	424	16 W14764	Human soluble kit
37	1233	54.9	424	16 W14765	Human soluble kit
38	1233	54.9	435	13 R26530	Sequence of one ch
39	1233	54.9	437	18 W10552	Alpha-1-acid glyco
40	1233	54.9	442	18 W10550	IgG1 polypeptide.
41	1233	54.9	445	20 Y24153	Bovine LOX-1 extra
42	1233	54.9	446	17 W05829	Humanised LD10 ant
43	1233	54.9	447	20 Y31669	Human IgG1 chain C
44	1233	54.9	449	14 R43339	Completely humanis
45	1233	54.9	449	19 W49816	Amino acid sequenc

ALIGNMENTS

RESULT 1	
W48650	
ID W48650 standard; Protein; 652 AA.	
XX	
AC W48650;	
XX	
DT 04-AUG-1998 (first entry)	
XX	
DE Heavy chain of hmaB425 fused to TNF alpha.	
XX	
KW Antibody-cytokine fusion protein; tricistronic vector; chimeric;	
KW TNF alpha; IL-2; IRES; internal ribosome entry site.	
XX	
OS Synthetic.	
OS Homo sapiens.	
XX	
FH Key	Location/Qualifiers
FT Region	1..494
FT	/note= "Heavy chain of human mAb 425"
FT Region	495..652
FT	/note= "TNF alpha"
XX	
PN WO9811241-A1.	
XX	
PD 19-MAR-1998.	
XX	
PF 02-SEP-1997;	97WO-EP04765.
XX	
PR 30-SEP-1996;	96EP-0115635.
PR	16-SEP-1996;
XX	96EP-0114820.
XX	
PA (MERE) MERCK PATENT GMBH.	
XX	
PI Bruemmer W, Burge C, Dunker R, Hauser H, Mielke C;	
PI Rieke E, Von Hoegen I, Welge T;	

XX WPI; 1998-207400/18.
 DR N-PSDB; V18096.
 XX
 XX Oligo:cistronic expression vector - useful for production of, e.g.
 PT MAb425/TNF- α or MAb425/IL-2 antibody fusion protein
 PT
 XX Disclosure; Fig 15; 89pp; English.
 PS
 XX The present sequence represents a fusion protein comprising of TNF
 CC alpha fused to the C-terminus of the heavy chain of the human
 CC monoclonal antibody 425 (hmb425). The hmb425 has specificity for
 CC the human EGF receptor. The invention claims for a new pMCDHAP
 CC tricistronic vector (V18096) for the expression of an antibody-cytokine
 CC fusion protein, hmb425-TNF alpha. The TNF alpha sequence can be
 CC substituted by the IL-2 sequence. The vector also contains a strong
 CC promoter/enhancer unit, a selection marker gene and at least two
 CC poliovirus derived internal ribosomal entry site (IRES) sequences. The
 CC vector can be expressed in mammalian host cells for the production of
 CC heteromeric fusion proteins. This expression system is claimed to
 CC produce the heteromeric proteins in high yields.
 XX
 SQ Sequence 652 AA;

Query Match 55.2%; Score 1239.5; DB 19; Length 652;
 Best Local Similarity 65.8%; Pred. No. 9.4e-76;
 Matches 250; Conservative 19; Mismatches 54; Indels 57; Gaps 6;
 QY 2 DKHTTCCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
 Db 268 dkthtccpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkfnwyvd 327
 QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
 Db 328 gvevhnaaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiektiskak 387
 QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLIVKGFPYSDIAVEWESNGQPENNYKTPPPVLD 181
 Db 388 gqprepqvtytlppsrdeitknqvslctclvkgfypsdiavewesngqpennyktppvlds 447
 QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 241
 Db 448 dgsfflyskltvdksrwqgnvfscvsmhealhnhytqkslsispgknvrssrrtpskdp 507
 QY 242 TSHQLLCKDKPPGYTLKQHTAKWKTCAPCPDHYITDTSWHTSDCLYCSPVCK-ELQYV 300
 Db 508 vah-----vvanpqaegqlgl 524
 QY 301 KOECNR-----THNRVCECKEGRYL-ETEFCLKHRSCPPGGVVOAGTPERNWYCKR 351
 Db 525 nrranallangvelrdnglvvpseglyliysqvifkgggc-psthvllthtsrlavsyg 583
 QY 352 CPDGFFSNETSSKAPCRKHT 371
 Db 584 tkvnlls---aikspcqrct 600
 RESULT 2
 W83963
 ID W83963 standard; Protein; 374 AA.
 XX
 AC W83963;
 XX
 DT 05-FEB-1999 (first entry)
 XX
 DE Recombinant human metFc-OB protein variant.
 XX
 KW Recombinant; metFc-OB protein; Fc region; immunoglobulin; Ig; OB;
 KW obesity; human; adiposity; blood lipid; diabetes type II; insulin;
 KW hypoglycaemic; antihypertensive; diuretic; appetite suppressant;
 KW suspension; variant.
 XX

OS Homo sapiens.
 XX
 PN W09846257-A1.
 XX
 PD 22-OCT-1998.
 XX
 PF 16-APR-1998; 98WO-US07828.
 XX
 PR 14-APR-1998; 98US-0059467.
 PR 17-APR-1998; 97US-0843971.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Brems DN, French DL, Speed MA;
 DR WPI; 1998-594525/50.
 XX N-PSDB; V69686.
 PT Concentrated suspension of fusion of obesity protein with Fc
 PT immunoglobulin fragment - stable at physiological pH, used for e.g.
 PT reduction of weight and blood lipid levels, and for treatment of
 PT type II diabetes
 XX
 PS Claim 2; Fig 6A-C; 47pp; English.
 XX
 CC This represents a recombinant metFc-OB protein variant which consists of
 CC an Fc region of human immunoglobulin (Ig) fused to a human OB (obesity)
 CC protein. The invention provides a human OB protein suspension that
 CC contains at least 0.5 mg/ml of the human OB protein derivatised by
 CC attachment of the Fc region of an Ig to the N-terminus of OB, and has a
 CC pH 6-8. The suspensions are used to reduce weight, adiposity and blood
 CC lipid levels, to treat or prevent diabetes type II, and to increase lean
 CC mass and insulin sensitivity. They may be used in conjunction with
 CC insulin, hypoglycaemics, antihypertensives, diuretics, appetite
 CC suppressants etc. These suspensions are stable and active at
 CC physiological pH and are ready-for-use formulations that do not require
 CC freezing or freeze drying. As they are very concentrated, only small
 CC volumes are required and they provide a sustained-release effect, with
 CC increased potency and reduced frequency of injection.
 XX
 SQ Sequence 374 AA;
 Query Match 55.1%; Score 1238; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 6.1e-76;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
 Db 1 mdkthtccpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdshedpevkfnwyv 60
 QY 61 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 120
 Db 61 dgvevhnaaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiektiska 120
 QY 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLIVKGFPYSDIAVEWESNGQPENNYKTPPPVLD 180
 Db 121 kgqprepqvtytlppsrdeitknqvslctclvkgfypsdiavewesngqpennyktppvld 180
 QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 228
 Db 181 sdgsfflyskltvdksrwqgnvfscvsmhealhnhytqkslsispgk 228
 RESULT 3
 W49075
 ID W49075 standard; Protein; 374 AA.
 XX
 AC W49075;
 XX
 DT 18-NOV-1998 (first entry)
 XX
 DE Recombinant human MetFc-OB variant 2 protein.

XX Recombinant human MetFc-OB variant 2 protein; chimeric; immunoglobulin;
 KW high blood lipid level; arterial sclerosis; stroke; Fc-OB fusion protein;
 KW diabetes.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX Location/Qualifiers
 FT Protein
 FT 2..374
 FT /note= "Recombinant human Fc-OB variant 2 protein"
 FT Region
 FT 229..374
 FT /note= "Human OB protein"
 XX
 XX WO9828427-A1.
 XX
 XX 02-JUL-1998.
 XX
 XX 11-DEC-1997; 97WO-US23183.
 XX
 XX 20-DEC-1996; 96US-0770973.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 XX Hecht RI, Mann MB;
 XX
 XX WPI; 1998-377658/32.
 XX N-PSDB; V32902.
 XX
 XX New fusion proteins of OB and Fc - used for treating e.g. excess
 PT weight, diabetes, arterial sclerosis, arterial plaque, high blood
 PT lipid level, gall stones or stroke
 XX
 XX Claim 2; Fig 5A-5C; 107pp; English.
 XX
 CC The present sequence represents a recombinant human MetFc-OB variant 2
 CC fusion protein having a 5 residue deletion of residues 2-6 of the
 CC wild-type Fc-OB protein sequence shown in W45073. The invention provides
 CC Fc-OB fusion proteins whereby the Fc region of an immunoglobulin or its
 CC analogue is linked, either directly or indirectly using a linker, to the
 CC N-terminus of an OB protein or its analogue. The Fc-OB fusion proteins
 CC are claimed to demonstrate increased stability and clearance rate and
 CC decreased degradation as compared to OB protein or a fusion of Fc to
 CC the C-terminus of the OB protein. These Fc-OB fusion proteins are also
 CC claimed to be useful for treating excess weight in an individual or
 CC animal or for treating co-morbidities associated with excess fat such as
 CC diabetes, high blood lipid level, arterial sclerosis and stroke.
 XX
 SQ Sequence 374 AA;

Query Match 55.1%; Score 1238; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 6.1e-76;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDKTHTCPCPAPALGGPSVFLPPPKDGLMISRTPEVTCVVVDSHEDPEVKFNWY 60
 DB 1 mdkthtcpcpapalggpsvflpppkdglmistrpevtcvcvvdshedpevkfnwv 60
 QY 61 DGEVHNKTKPREEQNSTYRVVSVTLVLDHQLNGKEYKCKVSNKALPAPIETISKA 120
 DB 61 dgevhnaktpreeqnystyrvvsvtlvldhqlngkeykckvsnkalpapietiska 120
 QY 121 KGQPREQVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLD 180
 DB 121 kgqpreqvyltppsrdeltknqvslltclvkgfypsdiavwesngqpennnykttppvld 180
 QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
 DB 181 sdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispgk 228
 RESULT 4

Y01372
 ID Y01372 standard; Protein; 235 AA.
 XX Y01372;
 AC Y01372;
 XX
 DT 04-JUN-1999 (first entry)
 XX
 DE Amino acid sequence of Fc fragment of human IgG.
 XX
 KW FcRn; binding; epithelial cell; immune system; modulation; antigen;
 KW pathogen; autoimmune disease; allergen; tumour; therapeutic; cytokine;
 KW chemotherapy agent; interferon; insulin; human growth hormone; fertility;
 KW drug; calcitonin; steroid; immunity; mucosal; AIDS; hepatitis; Fc; human;
 KW immunoglobulin; IgG.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 230
 FT /note= "encoded by TGA"
 XX
 XX WO9904813-A1.
 XX
 XX 04-FEB-1999.
 XX
 XX 24-JUL-1998; 98WO-US15395.
 XX
 XX 24-JUL-1997; 97US-0899856.
 XX
 XX (BGHM) BRIGHAM & WOMENS HOSPITAL.
 XX (UYBR-) UNIV BRANDEIS.
 XX
 XX Blumgerg RS, Lencer WI, Simister NE;
 XX
 XX WPI; 1999-153297/13.
 XX N-PSDB; X27800.
 XX
 XX Delivering therapeutics, particularly antigens to epithelial cells -
 XX comprises use of a conjugate of the therapeutic and an FcRn binding
 XX partner
 XX
 XX Disclosure; Fig 1; 74pp; English.
 XX
 CC The invention relates to methods of delivering molecules to a mammal by
 CC administering a conjugate of a therapeutic or a bioactive substance,
 CC and an FcRn binding partner targeted to epithelial cells expressing FcRn.
 CC This can be used for modulating a mammalian immune system wherein, the
 CC bioactive substance is an antigen characteristic of a pathogen of an
 CC autoimmune disease, an allergen, or a tumour. The delivering method is
 CC used to deliver therapeutics, including chemotherapy agents, cytokines
 CC including interferon, hormones, including insulin and human growth
 CC hormone, fertility drugs, calcitonin, calcitriol and other bioactive
 CC steroids to intestinal, mucosal and lung epithelium. The method is also
 CC used for trans epithelial delivery of antigens to provoke tolerance and
 CC immunity. The method provides an immunisation that specifically targets
 CC the mucosal surfaces, and does not have the risks associated with
 CC injections, including needle transmission of AIDS or hepatitis, of prior
 CC art immunisation methods. The present sequence represents the amino acid
 CC sequence of an Fc fragment of human immunoglobulin IgG.
 XX
 SQ Sequence 235 AA;

Query Match 55.1%; Score 1237; DB 20; Length 235;
 Best Local Similarity 97.4%; Pred. No. 4.1e-76;
 Matches 228; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MDKTHTCPCPAPALGGPSVFLPPPKDGLMISRTPEVTCVVVDSHEDPEVKFNWY 60
 DB 2 vdkthtcpcpapalggpsvflpppkdglmistrpevtcvcvvdshedpevkfnwv 61
 QY 61 DGEVHNKTKPREEQNSTYRVVSVTLVLDHQLNGKEYKCKVSNKALPAPIETISKA 120
 DB 61 dgevhnaktpreeqnystyrvvsvtlvldhqlngkeykckvsnkalpapietiska 120

Db 62 dgvevhnaktkpreeqnystyrsvsvltvlhqdwlngkeyckvsnkalpapietkiska 121
 QY 121 KGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAVWESNGQPNENYKTTTPPVLD 180
 Db 122 kgprepqvtytlppsrdeltnqvsitclvkgfypsdiavwesngqpennkyttppvld 181
 QY 181 SDGSFELYSLTVKSRWQGVNFCVSMHEALHNYTKSLSPGKETPPPK 234
 Db 182 sdgsfelyskltvdkrsrqggnvfscvsmhealhnytkslspgkvrrpr 235

RESULT 5
 W26649
 ID W26649 standard; Protein; 651 AA.
 AC W26649;
 XX
 DT 12-FEB-1998 (first entry)
 DE Chimeric receptor hCTM01/G1/zeta.

Cell activation: chimeric receptor; DNA delivery; CTM01; scFv;
 IgG1; T cell receptor zeta chain; cell proliferation; cytokine;
 inflammation; effector; cell differentiation; antibody secretion;
 phagocytosis; tumour infiltration; adhesion; infection; cancer;
 allergy; rheumatoid arthritis; osteoarthritis; asthma; eczema;
 inflammatory bowel disease; cystic fibrosis; sickle cell anaemia;
 psoriasis; multiple sclerosis; transplant rejection; diabetes;
 graft versus host disease; human; therapy.

Chimeric - Homo sapiens.
 OS
 XX
 XX
 PN WO9723613-A2.
 XX
 PD 03-JUL-1997.
 XX
 PF 23-DEC-1996; 96WO-GB03209.
 XX
 PR 21-DEC-1995; 95GB-0026131.
 XX
 XX (CLLT) CELLTech THERAPEUTICS LTD.

PI Bebbington CR, Finney HM, Lawson ADG, Weir ANC.
 XX
 XX WPI; 1997-351052/32.
 DR N-PSDB; T90512.
 XX
 PT New DNA systems for activating cells - comprising DNA coding for a
 PT chimeric receptor comprising 2 or more different cytoplasmic
 PT signalling components.
 XX
 PS Disclosure; Fig 7; 90pp; English.

CC This protein comprises a chimeric receptor consisting of an scFv
 CC engineered from anti-CD3 human antibody CTM01 linked to an
 CC extracellular spacer comprising the human IgG1 hinge, CH2 and CH3,
 CC linked to transmembrane and intracellular regions of the human T
 CC cell receptor zeta chain. It can be expressed in host cells (e.g.
 CC Jurkat) using a chimeric receptor gene (see T90512) constructed
 CC from DNA cassettes encoding each receptor component. In a claimed
 CC cell activation process an effector cell is transformed with DNA
 CC encoding a chimeric receptor containing 2 or more different
 CC cytoplasmic signalling components. Also claimed is use of DNA
 CC encoding a recombinant chimeric receptor in a DNA delivery system.
 CC The DNA delivery systems can be used for the activation of cells to
 CC provide e.g. an increase in cell proliferation, expression of
 CC cytokines with e.g. pro- or anti-inflammatory responses, stimulation
 CC of cytolytic activity, differentiation or other effector functions,
 CC antibody secretion, phagocytosis, tumour infiltration and/or
 CC increased adhesion. They can be used in the treatment of e.g.
 CC infectious disease, inflammatory disease, cancer, allergic/atopic
 CC disease, congenital disease, dermatologic disease, neurologic
 CC disease, transplants and metabolic/idiopathic disease (claimed).

CC In particular, they can be used in the treatment of rheumatoid
 CC arthritis, osteoarthritis, inflammatory bowel disease, asthma,
 CC eczema, cystic fibrosis, sickle cell anaemia, psoriasis, multiple
 CC sclerosis, organ or tissue transplant rejection, graft-versus-host
 CC disease or diabetes (claimed).
 XX
 SQ Sequence 651 AA;

Query Match 55.1%; Score 1236.5; DB 18; Length 651;
 Best Local Similarity 62.6%; Pred. No. 1.5e-75;
 Matches 258; Conservative 18; Mismatches 61; Indels 75; Gaps 9;

QY 2 DKTHTCPCPAPELLGSPVFLFPKPKDITLMISRTPEVTCVVVDSHEDPCKVNWVD 61
 Db 288 dkthtcpcppapellgspvflfpkpkdtlmisrtpevtcvvdsvedpckvfnvyd 347
 QY 62 GVEVHNAKTPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 121
 Db 348 gvevhnaktkpreeqnystyrsvsvltvlhqdwlngkeyckvsnkalpapietkiskak 407
 QY 122 GQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAVWESNGQPNENYKTTTPVLD 181
 Db 408 gqprepqvtytlppsrdeemtknqvsitclvkgfypsdiavwesngqpennkyttppvlds 467
 QY 182 DGSFELYSLTVKSRWQGVNFCVSMHEALHNYTKSLSPGKETPPPKLYLHDEE 241
 Db 468 dgsfelyskltvdkrsrqggnvfscvsmhealhnytkslspgk--ldpk----- 518
 QY 242 TSHOLLCKDKPPGYTLKQHTAKWKTCAPCDPHYTDWSHTSDCLYCSVPKELQYVK 301
 Db 519 -----lc-----ylldgilfygviltalflr 540
 QY 302 -----QECNPTHNVCECKEGRYLEIEFCLKHKRSCPPGFGVQAGTPERTIV 348
 Db 541 vkfsrsadapayqggnqlyn---elnlgrreeydvldkrrgrdpemg----gkpr--- 590
 QY 349 CKRCPDGFSSNETSSKAPCRKHTNCVFGLLLTOKGNATHDNCISGNSSTQ 400
 Db 591 -knpgqgly-nelqdkmaeayseigmgerrrkg---hdglyqglstak 637

RESULT 6
 W26650
 ID W26650 standard; Protein; 692 AA.
 XX
 AC W26650;
 XX
 DT 12-FEB-1998 (first entry)
 XX
 DE Chimeric receptor hCTM01/G1/zeta-CD28.
 XX
 KW Cell activation; chimeric receptor; DNA delivery; CTM01; scFv;
 KW IgG1; CD28; T cell receptor zeta chain; proliferation; cytokine;
 KW inflammation; effector; cell differentiation; antibody secretion;
 KW phagocytosis; tumour infiltration; adhesion; infection; cancer;
 KW allergy; rheumatoid arthritis; osteoarthritis; asthma; eczema;
 KW inflammatory bowel disease; cystic fibrosis; sickle cell anaemia;
 KW psoriasis; multiple sclerosis; transplant rejection; diabetes;
 KW graft versus host disease; human; therapy.
 XX
 OS Chimeric - Homo sapiens.
 XX
 PN WO9723613-A2.
 XX
 PD 03-JUL-1997.
 XX
 PF 23-DEC-1996; 96WO-GB03209.
 XX
 PR 21-DEC-1995; 95GB-0026131.
 XX
 XX (CLLT) CELLTech THERAPEUTICS LTD.
 XX

PI Bebbington CR, Finney HM, Lawson ADG, Weir ANC;
 XX WPI; 1997-351052/32.
 DR N-PSDB; T90513.
 XX
 PT New DNA systems for activating cells - comprising DNA coding for a
 PT chimeric receptor comprising 2 or more different cytoplasmic
 PT signalling components.
 XX
 PS Disclosure; Fig 8; 90pp; English.
 XX
 CC This protein comprises a chimeric receptor consisting of an scFv
 CC engineered from anti-CD3 human antibody CTMO1 linked to an
 CC extracellular spacer comprising the Human IgG1 hinge, CH2 and CH3,
 CC linked to the transmembrane and intracellular components of the
 CC human T cell receptor zeta chain, fused to the intracellular region
 CC of human CD28. It can be expressed in host cells (e.g. Jurkat)
 CC using a chimeric receptor gene (see T90513) constructed from DNA
 CC cassettes encoding each component of the receptor. In a claimed
 CC cell activation process an effector cell is transformed with DNA
 CC encoding a chimeric receptor containing 2 or more different
 CC cytoplasmic signalling components. Also claimed is use of DNA
 CC encoding a recombinant chimeric receptor in a DNA delivery system.
 CC The DNA delivery systems can be used for the activation of cells to
 CC provide e.g. an increase in cell proliferation, expression of
 CC cytokines with e.g. pro- or anti-inflammatory responses, stimulation
 CC of cytolytic activity, differentiation or other effector functions;
 CC antibody secretion, phagocytosis, tumour infiltration and/or
 CC increased adhesion. They can be used in the treatment of e.g.
 CC infectious disease, inflammatory disease, cancer, allergic/atopic
 CC disease, congenital disease, dermatologic disease, neurologic
 CC disease, transplants and metabolic/idiopathic disease (claimed).
 CC In particular, they can be used in the treatment of rheumatoid
 CC arthritis, osteoarthritis, inflammatory bowel disease, asthma,
 CC eczema, cystic fibrosis, sickle cell anaemia, psoriasis, multiple
 CC sclerosis, organ or tissue transplant rejection, graft-versus-host
 CC disease or diabetes (claimed).
 XX
 SQ Sequence 692 AA;

Query Match 55.1%; Score 1236.5; DB 18; Length 692;
 Best Local Similarity 62.6%; Pred. No. 1.6e-75;
 Matches 258; Conservative 18; Mismatches 61; Indels 75; Gaps 9;
 QY 2 DKTHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
 Db 288 dkthtcpcpapelggpsvflppkpkdtlmisrtpetvctvvdvshedpevkfnwyvd 347
 QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVTLHQDWLNGKEYCKVSNKALPAPIETKISKAK 121
 Db 348 gvevhnaktprreeqynstyrvvsvltvltlhbqdwlngkeyckvsnkalpapietkiskak 407
 QY 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD 181
 Db 408 gqpreqvtytlppsrdeetknqvsltcclvkgyfypsdiavwesngqpennnykttppvlds 467
 QY 182 DGSFELYSKLTVDKSRWQGVNVCSCVWHEALHNHYTQKSLSLSPGKETPPPKYLVHDEE 241
 Db 468 dgsfelyskltvdksrwqgvnvcscvwmhealnhytqkslsispgk--ldpk----- 518
 QY 242 TSHQLLCKPCPGGYTLKHOHTAKWKTVCAPCPDHYHYTDSWHTSDECLYCSVPCKELQYVK 301
 Db 519 ----lc-----yldglifigviltalflr 540
 QY 302 -----QECNRTNIRVCECKEGRYLEIEFCLKHRSCPPPGFVVQAGTPERTVY 348
 Db 541 vkfsrsadapayqggqqllyn---elnlgreeydvldkrgrdpeng----gkpr--- 590
 QY 349 CKRCPDGFNFSKAPCRKHNCSVFGLLLTQKGNATHDNCISGNSSTQ 400
 Db 591 -knpqegly-neiqdkmaeayseigmkerrrgk---hdglyqglstatk 637

RESULT 7
 R27163
 ID R27163 standard; Protein; 347 AA.
 XX
 AC R27163;
 XX
 DT 20-MAY-1998 (first entry)
 XX
 DE CD2 binding LFA-3-Ig fusion protein.
 XX
 KW lymphocyte associated antigen-3; T-lymphocyte accessory molecule;
 KW deletion mutant; CD2 binding site; immunomodulator; immunoglobulin;
 KW preLFA3TIP; LFA-3(92)IgG; pSAB152; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT peptide 1..28
 FT /label= LFA-3_signal
 FT region 29..120
 FT /label= LFA-3_aminoacids_1-92
 FT region 121..130
 FT /label= IgG1_hinge
 FT domain 131..239
 FT /label= IgG1_CH2
 FT domain 240..347
 FT /label= IgG1_CH3
 XX
 PN EP503648-A.
 XX
 PD 16-SEP-1992.
 XX
 PF 12-MAR-1992; 92EP-0104320.
 XX
 PR 12-MAR-1991; 91US-0667971.
 PR 07-OCT-1991; 91US-0770967.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Miller GT, Rosa MD, Wallner BP;
 XX
 DR WPI; 1992-309760/38.
 DR N-PSDB; Q28684.
 XX
 PT CD2-binding domain of lymphocyte function associated antigen-3
 PT and DNA - for diagnosing and treating inflammation and
 PT autoimmune diseases, e.g. systemic lupus erythematosus and
 PT rheumatoid arthritis
 PS
 PS Claim 13; Fig 12; 85pp; English.
 XX
 CC The plasmid pSAB152 contains the DNA sequence encoding the LFA-3
 CC signal sequence, the amino terminal 92 amino acids of mature LFA-3,
 CC ten amino acids of the hinge region of IgG1 and the CH2 and CH3
 CC constant domains of IgG1 (see Q28678-9 and Q28681-2 for details of
 CC the construction of pSAB152). A NotI fragment containing the coding
 CC sequence of pSAB152 was used in the construction of expression
 CC vector pMDR(92)ig-3 which can be stably maintained in CHO cells to
 CC achieve continuous expression of LFA3TIP. The fusion protein can
 CC bind to CD2 and inhibit T cell activation, making it useful to
 CC treat acute and chronic inflammation, autoimmune disease and
 CC in immunomodulation.
 XX
 SQ Sequence 347 AA;
 Query Match 54.9%; Score 1234; DB 13; Length 347;
 Best Local Similarity 59.6%; Pred. No. 1e-75;
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDKTHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
 :|||||

XX

PS Disclosure; Page 62-63; 69pp; English.

XX The present sequence represents a splice variant of a rabbit transforming
 CC growth factor-beta (TGF-beta) type II receptor fused to the Fc portion
 CC of human IgG1. The fusion protein has higher affinity for TGF-beta than
 CC fusion proteins comprising the non-variant form of the receptor. The
 CC fusion protein contains soluble TGF-beta receptor constructs that are
 CC devoid of a transmembrane region (and are secreted from the cell) but
 CC retain the ability to bind TGF-beta. The protein competitively inhibits
 CC binding of TGF-beta to cellular receptors and/or forms an inactive
 CC complex with TGF-beta. The protein is used to reduce levels of TGF-beta,
 CC for treatment of arthritic conditions associated with overexpression
 CC of TGF-beta, especially fibroproliferative diseases, e.g. renal,
 CC intra-ocular or pulmonary fibrosis; diabetic nephropathy;
 CC glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;
 CC collagen vascular disease, e.g. systemic sclerosis, polymyositis,
 CC scleroderma, dermatomyositis or systemic lupus erythematosus; and
 CC fibrosis associated with restenosis. It is also used for treating
 CC wounds, to prevent overproduction of connective tissue and so prevent
 CC adhesions or scarring, and to prevent post-radiation fibrosis (by
 CC administration to patients about to undergo radiation therapy).

XX Sequence 388 AA;

Query Match 54.9%; Score 1234; DB 21; Length 388;
 Best Local Similarity 99.6%; Pred. No. 1.2e-75;
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPCPAPPELLGGPSVFLFPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
 :|||||
 Db 161 vdkthtcpcpapellggpsvflfppkdtlmsrtpevtcvvvdvshedpevkfnwyv 220

Qy 61 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 120
 :|||||
 Db 221 dgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektiska 280

Qy 121 KGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 180
 :|||||
 Db 281 kgqprepqvyltppsrdeitknqsltcivkgfypsdiavewesngqpennnykttppvld 340

Qy 181 SDGSFFLYSKLTVDKSRQOQGVNFSCSYVMHEALHNHYTQKSLSLSPGK 228
 :|||||
 Db 341 sdgsfflyskltdvksrqoqgvnfvscsvmhealhnhytqkslsispgk 388

RESULT 12
 Y54064
 ID Y54064 standard; Protein; 388 AA.
 AC Y54064;
 XX
 XX 27-MAR-2000 (first entry)
 XX
 XX Amino acid sequence for TGF-beta type II receptor variant/IgG1 fusion.
 DE
 XX Splice variant; human; transforming growth factor-beta; TGF-beta;
 KW type II receptor; Fc portion; human IgG1; fusion protein; arthritis;
 KW fibroproliferative disease; renal; intra-ocular; pulmonary; fibrosis;
 KW diabetic nephropathy; glomerulonephritis; collagen vascular disease;
 KW proliferative vitreoretinopathy; myelofibrosis; systemic sclerosis;
 KW polymyositis; scleroderma; dermatomyositis; systemic lupus erythematosus;
 KW restenosis; wound; connective tissue production; adhesion; scarring;
 KW post-radiation fibrosis.
 XX
 XX Synthetic.
 OS
 OS Homo sapiens.
 XX
 XX WO9965948-A1.
 XX
 XX 23-DEC-1999.
 PD
 XX 16-JUN-1999; 99WO-US13629.

XX 16-JUN-1998; 98US-0089452.
 PR (BIOJ) BIOGEN INC.
 PA
 XX Koteliansky V, Gotwals P, Cate R, Sanicola-Nadel M;
 PI WPI; 2000-106083/09.
 XX N-PSDB; 245253, 245254.
 DR
 DR New fusion protein of a splice variant of transforming growth
 FT factor-beta receptor, for inhibiting the growth factor, e.g. in
 FT treatment of fibrosis
 PT
 PS Disclosure; Page 64; 69pp; English.

XX The present sequence represents a splice variant of a human transforming
 CC growth factor-beta (TGF-beta) type II receptor fused to the Fc portion
 CC of human IgG1. The fusion protein has higher affinity for TGF-beta than
 CC fusion proteins comprising the non-variant form of the receptor. The
 CC fusion protein contains soluble TGF-beta receptor constructs that are
 CC devoid of a transmembrane region (and are secreted from the cell) but
 CC retain the ability to bind TGF-beta. The protein competitively inhibits
 CC binding of TGF-beta to cellular receptors and/or forms an inactive
 CC complex with TGF-beta. The protein is used to reduce levels of TGF-beta,
 CC for treatment of arthritic conditions associated with overexpression
 CC of TGF-beta, especially fibroproliferative diseases, e.g. renal,
 CC intra-ocular or pulmonary fibrosis; diabetic nephropathy;
 CC glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;
 CC collagen vascular disease, e.g. systemic sclerosis, polymyositis,
 CC scleroderma, dermatomyositis or systemic lupus erythematosus; and
 CC fibrosis associated with restenosis. It is also used for treating
 CC wounds, to prevent overproduction of connective tissue and so prevent
 CC adhesions or scarring, and to prevent post-radiation fibrosis (by
 CC administration to patients about to undergo radiation therapy).

XX Sequence 388 AA;

Query Match 54.9%; Score 1234; DB 21; Length 388;
 Best Local Similarity 99.6%; Pred. No. 1.2e-75;
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPCPAPPELLGGPSVFLFPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
 :|||||
 Db 161 vdkthtcpcpapellggpsvflfppkdtlmsrtpevtcvvvdvshedpevkfnwyv 220

Qy 61 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 120
 :|||||
 Db 221 dgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektiska 280

Qy 121 KGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 180
 :|||||
 Db 281 kgqprepqvyltppsrdeitknqsltcivkgfypsdiavewesngqpennnykttppvld 340

Qy 181 SDGSFFLYSKLTVDKSRQOQGVNFSCSYVMHEALHNHYTQKSLSLSPGK 228
 :|||||
 Db 341 sdgsfflyskltdvksrqoqgvnfvscsvmhealhnhytqkslsispgk 388

RESULT 13
 Y70867
 ID Y70867 standard; Protein; 399 AA.
 AC Y70867;
 XX
 XX 31-JUL-2000 (first entry)
 XX
 XX Human interferon-beta-1a/mouse IgG2a Fc fusion protein.
 DE
 XX Human; interferon-beta-1a; IFN-beta-1a; immunoglobulin; fusion protein;
 KW angiogenesis; antisclerotic; antiinflammatory; immunosuppressive; cancer;
 KW cytostatic; virucide; hepatotropic; antiangiogenic; treatment; fibrosis;

KW multiple sclerosis; inflammatory disease; autoimmune disease; hepatitis;
 KW viral infection; neovascularisation; mouse; IgG2a Fc domain.
 OS Chimeric - Homo sapiens.
 OS Chimeric - Mus sp.

XX Key Location/Qualifiers
 FH 1..166
 FT /label= natural_human_IFN-beta-la
 FT 167..171
 FT /label= Enterokinase_linker
 FT 172..399
 FT /label= Mouse_IgG2a_Fc_portion
 FT /note= "comprises the hinge region, CH2 and CH3
 FT constant domains of mouse Ig"

WO200023472-A2.

XX 27-APR-2000.
 XX 15-OCT-1999; 99WO-US24200.
 XX 16-OCT-1998; 98US-0104491.
 PR 16-FEB-1999; 99US-0120237.
 XX (BIOJ) BIOGEN INC.

XX Whitty A, Runkel L, Brickelmaier M, Hochman P;
 XX WPI; 2000-339654/29.
 DR N-PSDB; D00167.
 XX Fusion proteins comprising interferon-beta-la useful for inhibiting
 PT angiogenesis -
 XX Example 2; Fig 2; 82pp; English.

XX The patent discloses fusion proteins comprising glycosylated
 CC interferon-beta (IFN-beta) especially IFN-beta-la, linker groups and
 CC non-IFN-beta proteins, especially an immunoglobulin (Ig) protein.
 CC The present sequence is a fusion protein that consists
 CC of mature human IFN-beta-la and mouse IgG2a Fc domain separated by an
 CC enterokinase linker. The fusion protein is useful for
 CC inhibiting angiogenesis in a patient. It may also be used to treat
 CC multiple sclerosis, fibrosis, inflammatory and autoimmune diseases,
 CC cancers, hepatitis and viral infection characterised by
 CC neovascularisation.

XX Sequence 399 AA;

Query Match 54.9%; Score 1234; DB 21; Length 399;
 Best Local Similarity 99.6%; Pred. No. 1.2e-75;
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPCAPPELLGSPSVFLPPPKDGLMISRTPEVTGVVDVSHEDPEVKENMYV 60
 Db :|||||
 172 vdkthcpcpcapellgppsvflppkpkdglmsrtpevtcvvdvshedpevkfnvyv 231
 QY 61 DGVEVHNKTKPREQYNSTYRYVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 120
 Db |||||
 232 dgvevhnaktpreeqynstyrsvvltvlhqdwlngkeyckvsnkalpapiektiska 291
 QY 121 KGQPREQVYTLPSRDELKQVSLTCLVKGFPSPDIWESNGQENNYKTPPVLD 180
 Db |||||
 292 kgqpreqvylpsrdeltkngvsltclvkgfypsdiawesngqpennnykttppvld 351

QY 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKQSLSLSPCK 228
 Db |||||
 352 sdgsfflyskltvdksrwqgnvfscvmhealhnhytkqslslspck 399

RESULT 14

Y70868
 ID Y70868 standard; Protein; 418 AA.
 AC Y70868;
 XX 31-JUL-2000 (first entry)
 DE Human interferon-beta-la G162C-IgG1 Fc direct fusion protein.
 XX Human; interferon-beta-la; IFN-beta-la; immunoglobulin; fusion protein;
 KW angiogenesis; antisclerotic; antiinflammatory; immunosuppressive; cancer;
 KW cytostatic; virucide; hepatotropic; antiangiogenic; treatment; fibrosis;
 KW multiple sclerosis; inflammatory disease; autoimmune disease; hepatitis;
 KW viral infection; neovascularisation; IgG1 Fc domain; VCAM-1; mutant;
 KW vascular cell adhesion molecule-1.
 XX Homo sapiens.
 OS Synthetic.
 OS Key Location/Qualifiers
 FH 1..24
 FT /label= VCAM-1_signal_sequence
 FT /note= "Vascular cell adhesion molecule-1"
 FT 25..190
 FT /label= natural_human_IFN-beta-la
 FT Misc-difference 186
 FT /note= "Wild type IFN-beta-la Gly at position 162 is
 FT substituted by Cys"
 FT 192..418
 FT /label= Human_IgG1_Fc_portion
 FT /note= "comprises hinge, CH2 and CH3 domains of
 FT IgG1 heavy chain"
 XX WO200023472-A2.
 XX 27-APR-2000.
 XX 15-OCT-1999; 99WO-US24200.
 XX 16-OCT-1998; 98US-0104491.
 PR 16-FEB-1999; 99US-0120237.
 XX (BIOJ) BIOGEN INC.
 XX Whitty A, Runkel L, Brickelmaier M, Hochman P;
 XX WPI; 2000-339654/29.
 DR N-PSDB; D00167.
 XX Fusion proteins comprising interferon-beta-la useful for inhibiting
 PT angiogenesis -
 XX Example 5; Fig 10; 82pp; English.

XX The patent discloses fusion proteins comprising glycosylated
 CC interferon-beta (IFN-beta) especially IFN-beta-la, linker groups and
 CC non-IFN-beta proteins, especially an immunoglobulin (Ig) protein.
 CC The present sequence is a fusion protein consisting of a vascular
 CC cell adhesion molecule-1 (VCAM-1) signal sequence, a modified human
 CC IFN-beta-la and human IgG1 Fc domain, which are directly fused without
 CC a linker sequence. The fusion protein is useful for
 CC inhibiting angiogenesis in a patient. It may also be used to treat
 CC multiple sclerosis, fibrosis, inflammatory and autoimmune diseases,
 CC cancers, hepatitis and viral infection characterised by
 CC neovascularisation.
 XX Sequence 418 AA;

Query Match 54.9%; Score 1234; DB 21; Length 418;
 Best Local Similarity 99.6%; Pred. No. 1.3e-75;
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
:|||||
Db 191 vdkthtcpcpapellggspsvflfppkpkdtlmisrtpevtcvvvdshedpevkfnwv 250
Qy 61 DGVEVHNAKTKPREEQYNSTYRVSVLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLD 120
:|||||
Db 251 dgvevhnaaktprreeqynstyrsvsylvlqhgdwlngkeyckvsnkalpapietiska 310
Qy 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLD 180
:|||||
Db 311 kgqprepqvylppsrdeltnqvsltcclvkgyfypsdiavwesngqpennnykttppvld 370
Qy 181 SDGSFFLYSKLTVDKSRWQQGNVSCSYMHEALHNHYTQKSLSLSPGK 228
:|||||
Db 371 sdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispgk 418

RESULT 15
Y70869
ID Y70869 standard; Protein; 423 AA.
XX AC Y70869;
XX DT 31-JUL-2000 (first entry)
XX DE Human interferon-beta-1a G162C/G4S linker/IgG1 Fc fusion protein.
XX KW Human; interferon-beta-1a; IFN-beta-1a; immunoglobulin; fusion protein;
KW angiogenesis; antisclerotic; antiinflammatory; immunosuppressive; cancer;
KW cytostatic; virucide; hepatotropic; antiangiogenic; treatment; fibrosis;
KW multiple sclerosis; inflammatory disease; autoimmune disease; hepatitis;
KW viral infection; neovascularisation; IgG1 Fc domain; VCAM-1; mutant;
KW vascular cell adhesion molecule-1; G4S linker.
XX OS Homo sapiens.
XX OS Synthetic.

XX FH Key Location/Qualifiers
FT Peptide 1..24
FT /label= VCAM-1_signal_sequence
FT /note= "Vascular cell adhesion molecule-1"
FT Protein 25..190
FT /label= natural_human_IFN-beta-1a
FT Misc-difference 186
FT /note= "Wild type IFN-beta-1a Gly at position 162 is
FT substituted by Cys"
FT Region 191..195
FT /label= G4S_linker
FT Region 197..423
FT /label= Human_IgG1_Fc_portion
FT /note= "comprises hinge, CH2 and CH3 domains of
FT IgG1 heavy chain"

XX WO2000023472-A2.
XX PD 27-APR-2000.
XX PF 15-OCT-1999; 99WO-US24200.
XX PR 16-OCT-1998; 98US-0104491.
XX PR 16-FEB-1999; 99US-0120237.
XX PA (BIOJ) BIOGEN INC.

XX PI Whitty A, Runkel L, Brickelmaier M, Hochman P;
XX WPI; 2000-339654/29.
XX DR N-PSDB; D00168.
XX PT Fusion proteins comprising interferon-beta-1a useful for inhibiting
XX angiogenesis -
XX PS Example 5; Fig 11; 82pp; English.

XX The patent discloses fusion proteins comprising glycosylated
CC interferon-beta (IFN-beta) especially IFN-beta-1a, linker groups and
CC non-IFN-beta proteins, especially an immunoglobulin (Ig) protein.
CC The present sequence is a fusion protein consisting of a vascular
CC cell adhesion molecule-1 (VCAM-1) signal sequence, a modified human
CC IFN-beta-1a which is linked to human IgG1 Fc domain via a G4S
CC linker sequence. The fusion protein is useful for
CC inhibiting angiogenesis in a patient. It may also be used to treat
CC multiple sclerosis, fibrosis, inflammatory and autoimmune diseases,
CC cancers, hepatitis and viral infection characterised by
CC neovascularisation.
XX SQ Sequence 423 AA;

Query Match 54.9%; Score 1234; DB 21; Length 423;
Best Local Similarity 99.6%; Pred. No. 1.3e-75;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
:|||||
Db 196 vdkthtcpcpapellggspsvflfppkpkdtlmisrtpevtcvvvdshedpevkfnwv 255
Qy 61 DGVEVHNAKTKPREEQYNSTYRVSVLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLD 120
:|||||
Db 256 dgvevhnaaktprreeqynstyrsvsylvlqhgdwlngkeyckvsnkalpapietiska 315
Qy 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLD 180
:|||||
Db 316 kgqprepqvylppsrdeltnqvsltcclvkgyfypsdiavwesngqpennnykttppvld 375
Qy 181 SDGSFFLYSKLTVDKSRWQQGNVSCSYMHEALHNHYTQKSLSLSPGK 228
:|||||
Db 376 sdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispgk 423

Search completed: March 1, 2001, 09:17:10
Job time: 228 sec